

TITLE	The three-dimensional solution structure of NAD1, a new floral defensin from Nicotiana glauca and its application to a homology model of the crop defense protein alfalfa					
JOURNAL	J. Mol. Biol. 325 (1), 175-188 (2003)					
MEDLINE	22361678					
PUBMED	12473460					
REFERENCE	2 (bases 1 to 601)					
AUTHORS	Lay,F.T., Brugliera,F. and Anderson,M.A.					
TITLE	Isolation and Properties of Floral Defensins from Ornamental Tobacco and Petunia					
JOURNAL	Plant Physiol. 131 (3), 1283-1293 (2003)					
PUBMED	12644678					
REFERENCE	3 (bases 1 to 601)					
AUTHORS	Lay,F.T. and Anderson,M.A.					
TITLE	Direct Submission					
JOURNAL	Submitted (06-May-2002) Biochemistry, La Trobe University, Bundoora, VIC 3086, Australia					
FEATURES	Location/Qualifiers					
source	1..601					
gene	/organism="Nicotiana glauca"					
	/mol_type="mRNA"					
	/db_xref="taxon:4087"					
CDS	1..601					
	/gene="Nad1"					
	61..378					
	/note="Nad1, gamma-thionin; possesses antifungal activity"; targeted to the vacuole"					
	/codon_start=1					
	/product="flower-specific defensin precursor"					
	/protein_id="AAW70999.1"					
	/translation="MARSDCEMAFLAMMLFVAIEVQARECKTSNPFGICITKPECRKAISKEFTDGHCSKLIRRLCTPKPCVFPEKMRTGAELIAEAKTLAALLDEEE"					
	MDN"					
	136..276					
	/gene="Nad1"					
	/product="flower-specific defensin"					
BASE COUNT	192 a 95 c 115 g 199 t					
ORIGIN						
Query Match	99.7%; Score 539.4; DB 8; Length 601;					
Best Local Similarity	99.8%; Pred. No.6.3e-102;					
Matches 540; Conservative	0; Mismatches 1; Indels 0; Gaps 0;					
Dn	1 ATGGCTGCGTCTGTGGCTTCATNGCATTTGCTATCTTGAGCAAGATGCTTTGTGGC 60					
Qy						
Dn	61 ATGGCTGCTGCTGTGGCTTCATNGCATTTGCTATCTTGAGCAAGATGCTTTGTGGC 120					
Qy						
Dn	61 TATGAGTGCAAGCTTAGAATGCACAAACAGAAGCACACATTCTGGAATATGCAAT 120					
Qy						
Dn	121 TATGAGTGCAAGCTTAGAATGCACAAACAGAAGCACACATTCTGGAATATGCAAT 180					
Qy						
Dn	121 ACCAACCACCCATNGCAAGAAAGTGTATACAGAGCAAATTTACTAGTATGCTATGAC 180					
Qy						
Dn	181 ACCAACCACCCATNGCAAGAAAGTGTATACAGAGCAAATTTACTAGTATGCTATGAC 240					
Qy						
Dn	181 AAAATCCTCAGAAAGTGCTATATTAAGCCATGTGTGTTGATGAGAAAGATGATTA 240					
Qy						
Dn	241 ACAGAGCGTAATTTTGGCTGAGGAAACAACCTTGGCTGACCTTTGCTTGAAGA 300					
Qy						
Dn	301 ACAGAGCGTAATTTTGGCTGAGGAAACAACCTTGGCTGACCTTTGCTTGAAGA 360					
Qy						
Dn	301 GAGATATGATATACTAATTTAGATTTAGAAATTTAAGATGACATACACATTAAT 360					
Qy						
Dn	361 GAGATATGATATACTAATTTAGATTTAGAAATTTAAGATGACATACACATTAAT 420					
Qy						
Dn	361 AAGTTTCTACCTTTCTTAAGAGTATGCTAATGTTGTTAATGGCTTTAGTAGC 420					
Qy						
Dn	421 AAGTTTCTACCTTTCTTAAGAGTATGCTAATGTTGTTAATGGCTTTAGTAGC 480					

Qy	421	CTTTATTACACTTAAATAGGTGGACCTTCAATCCTTTGGCAATCTTGCACTAAGT	480
Db	481	CTTTATTACACTTAAATAGGTGGACCTTCAATCCTTTGGCAATCTTGCACTAAGT	540
Qy	481	TTATTGTGACTTTTATGAATAAGCTTCTTAATGCTCTTGGTTAAAAA	540
Db	541	TTATTGTGACTTTTATGAATAAGCTTCTTAATGCTCTTGGTTAAAAA	600
Qy	541 A 541		
Db	601 A 601		

RESULT 2			
SEPTRNA		589 bp	RNA linear
LOCUS			PLN 03-DEC-1993
DEFINITION			Solanaceae mRNA for flower-specific thionin.
ACCESSION			Z11748 S41631
VERSION			Z11748.1 GI:21212
KEYWORDS			thionin.
SOURCE			Solanaceae
ORGANISM			Solanaceae
REFERENCE			Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Asteridae; Lamiales; Solanales.
AUTHORS			1 (bases 1 to 589)
TITLE			Gu,Q., Kawata,E.E., Morse,M.J., Wu,H.M. and Cheung,A.Y.
JOURNAL			A flower-specific cDNA encoding a novel thionin in tobacco
MEDLINE			Mol. Gen. Genet. 234 (1), 89-96 (1992)
PMID			92357021
REFERENCE			1495489
AUTHORS			2 (bases 1 to 589)
TITLE			Gu,Q.
JOURNAL			Direct Submission
FEATURES			Submitted (25-FEB-1992) Qing Gu, Department of Biology, Yale University, 165 Prospect St, New Haven, CT, 06511, USA
source			Location/Qualifiers
			1..589
			/organism="Solanaceae"
			/mol_type="mRNA"
			/db_xref="taxon:4070"
			/clone="49A"
			/issue_type="flower of tobacco"
			/dev_stage="flowering"
			49..366
			/codon_start=1
			/product="flower-specific thionin"
			/protein_id="CAA77806.1"
			/db_xref="GI:21213"
			/db_xref="SWISS-PROT:P32026"
			/translation="MARSLCMAFAALLMMFLFVAYVOARCKTESNFPGICITKPE
			CRKACISKEFTDGHCKSLRRCLCTKCPVPBKMLKIGAEFLVEAKTLAALLBEE
			MDN"

BASE COUNT	184 a	97 C	114 G	194 T
ORIGIN				
Query Match		93.2%	Score 504.2;	DB 8; Length 589;
Best Local Similarity		95.7%	Pred. No. 1.3e-94;	
Matches 518;		Conservative	0; Mismatches 23;	Indels 0; Gaps 0
Qy	1	ATGGCTGCTCTCTTGCTTCATGCGCAATTTGCTATCTTGCGAAGATGCTCTTGTGGC	60	
Db	49	ATGGCTGCTCTCTTGCTTCATGCGCAATTTGCTATCTTGCGAAGATGCTCTTGTGGC	108	
Qy	61	TATGAGTGCAAGCTTAGAATGCAAAACAGAAAGCAACATTCCTCGAATATGACAT	120	
Db	109	TATGAGTGCAAGCTTAGAATGCAAAACAGAAAGCAATACATTCCTCGAATATGACAT	168	
Qy	121	ACCAACCAACATGAGAAAGCTTGATCACTGAGAAATTTACTGATGCTATTGTAGC	180	
Db	169	ACCAACCAACATGAGAAAGCTTGATCACTGAGAAATTTACTGATGCTATTGTAGC	228	
Qy	181	AAATCTCAGAAAGTGCCTATGTACTAAGCAATGTGTGTTGATGAGAAATGACTAA	240	

Db 229 AAACCTCCTAGAGAGCTGCTATGACCTAAGCCTAGTGTGTGATGAGAAGATGATATAAA 288  
Qy 241 ACAGAGCTGATAATTTGGCTGAGAGAGCAAAACCTTGGCTGACCTTGGCTTGAAGAA 300  
Db 289 ACAGAGCTGATAATTTGGCTGAGAGAGCAAAACCTTGGCTGACCTTGGCTTGAAGAA 348  
Qy 301 GAGATAATGATTAATTAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 360  
Db 349 GAGATAATGATTAATTAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 408  
Qy 361 AAAGTTCTACCTTCTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 420  
Db 409 AAAGTTCTACCTTCTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 468  
Qy 421 CTTTATTAACCTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 480  
Db 469 CTTTATTAACCTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 528  
Qy 481 TTATTTGCTGCTTTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 540  
Db 529 TTATTTGCTGCTTTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 588  
Qy 541 A 541  
Db 589 A 589

RESULT 3  
AB005266 558 bp mRNA linear PLN 05-FEB-1999  
LOCUS Nicotiana excelsior mRNA for gamma-chionin, complete cds.  
DEFINITION AB005266  
ACCESSION AB005266.1 GI:2244704  
VERSION  
KEYWORDS gamma-chionin.  
SOURCE Nicotiana excelsior  
ORGANISM Nicotiana excelsior  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Nicotiana.  
REFERENCE  
1 (sites)  
Yamada, S., Komori, T. and Imaseki, H.  
CDNA cloning of gamma-chionin from Nicotiana excelsior (Accession  
No. AB005266) (PGR97-131)  
JOURNAL Plant Physiol. 115, 314 (1997)  
AUTHORS Yamada, S.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUN-1997) Shigehiro Yamada, Plant Breeding and  
Genetics Research Laboratory, Japan Tobacco Inc., 700 Higashibara,  
Iwata, Shizuoka 438, Japan  
(E-mail: Shigehiro.Yamada@pbgrl.jti.co.jp, Tel: 0538-32-7116,  
Fax: 0538-32-8700)

FEATURES  
source  
1..558  
Location/Qualifiers  
/organism="Nicotiana excelsior"  
/mol\_type="mRNA"  
/db\_xref="taxon:61185"  
1..558  
/gene="Nethio2"  
33..350  
/gene="Nethio2"  
/codon\_start=1  
/product="gamma-chionin"  
/protein\_id="BAA21114.1"  
/db\_xref="GI:2244705"  
/translation="MARSVCPMAFAIILAVLFEVAYDEAKDCKSTPFGICITKPP  
CRKACIKKFTDGHCKSLRCLCKPCVPEBKIKIKGATLAEATTLAAALREBI  
MDN"

BASE COUNT 170 a 94 c 108 g 186 t  
ORIGIN

Query Match

89.2%; Score 482.8; DB 8; Length 558;

Best Local Similarity 94.9%; Pred. No. 3.5e-90;  
Matches 499; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1 ATGGCTGGCTCCCTGGCTGCTTCAATGCTATCTTGGCAGAGATCTCTTTGTC 60  
Db 33 ATGGCTGGCTCCCTGGCTGCTTCAATGCTATCTTGGCAGAGATCTCTTTGTC 92  
Qy 61 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 93 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 152  
Qy 121 ACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180  
Db 153 ACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 212  
Qy 181 AAAATCCTAG 240  
Db 213 AAAATCCTAG 272  
Qy 241 ACAGAGCTGATAATTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
Db 273 ACAGAGCTGATAATTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 332  
Qy 301 GAGATAATGATTAATTAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 360  
Db 333 GAGATAATGATTAATTAAGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAG 392  
Qy 361 AAAGTTCTACCTTCTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 420  
Db 393 AAAGTTCTACCTTCTTAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 452  
Qy 421 CTTTATTAACCTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 480  
Db 453 CTTTATTAACCTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 512  
Qy 481 TTATTTGCTGCTTTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 526  
Db 513 TTATTTGCTGCTTTTAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 558

RESULT 4  
AB005250 549 bp mRNA linear PLN 19-JUN-1999  
LOCUS Nicotiana paniculata mRNA for gamma-chionin, complete cds.  
DEFINITION AB005250  
ACCESSION AB005250.1 GI:2251080  
VERSION  
KEYWORDS gamma-chionin.  
SOURCE Nicotiana paniculata  
ORGANISM Nicotiana paniculata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Nicotiana.  
REFERENCE  
1 (sites)  
Komori, T., Yamada, S. and Imaseki, H.  
A cDNA clone for gamma-chionin from Nicotiana paniculata (Accession  
No. AB005250) (PGR97-132)  
JOURNAL Plant Physiol. 115, 314 (1997)  
AUTHORS Komori, T.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUN-1997) Toshitsugu Komori, Plant Breeding and  
Genetics Research Laboratory, Japan Tobacco Inc., 700 Higashibara,  
Iwata, Shizuoka 438, Japan  
(E-mail: Toshitsugu.Komori@pbgrl.jti.co.jp, Tel: 0538-32-7116,  
Fax: 0538-32-8700)

FEATURES  
source  
1..549  
Location/Qualifiers  
/organism="Nicotiana paniculata"  
/mol\_type="mRNA"  
/db\_xref="taxon:62141"  
1..549  
/gene="NpThio1"  
48..368

gene  
CDS

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: January 28, 2004, 08:46:51 ; Search time 14 Seconds  
(without alignments)  
352.700 Million cell updates/sec

Title: US-10-072-809B-18

Perfect score: 552

Sequence: 1 MARSLCFMFAALARMFLVA.....LAEAKTLAALIEBIMDN 105

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt 41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	529	95.8	105	THGF_TOBAC
2	468.5	84.9	106	THG1_NICOT
3	144	26.1	73	THG1_PETIN
4	143.5	25.1	74	THG4_ARATH
5	138.5	24.9	77	P322_SOLTU
6	137.5	24.9	77	THG1_ARATH
7	137.5	24.9	77	THG3_ARATH
8	137	24.8	78	THGF_HELAN
9	135	24.5	74	DEP2_CAPAN
10	132	23.9	52	DEP2_CAPAN
11	127.5	23.1	77	THG2_ARATH
12	123	20.3	75	DEP1_CAPAN
13	112	20.3	80	DEP2_ARATH
14	110	19.9	74	SRP_SOYBN
15	108	19.4	80	APF1_BRANA
16	107	19.4	80	APF4_BRANA
17	106.5	19.3	79	THG1_MAIZE
18	105	19.0	80	APF3_RAPSA
19	105	18.7	80	APF3_ARATH
20	103	18.7	47	SAI1_SORBI
21	103	18.7	47	SAI1_SORBI
22	103	18.7	80	APF1_ARATH
23	103	18.6	80	APF2_RAPSA
24	102.5	18.6	79	APF3_BRANA
25	102	18.5	47	THG1_VICNA
26	101	18.3	47	FAH1_VICNA
27	101	18.3	47	FAH2_VICNA
28	99	17.9	47	THG HORVU
29	99	17.9	48	SIK2_SORBI
30	97	17.6	80	APF4_ARATH
31	94.5	17.1	74	DR92_PEA
32	94	17.0	47	THG2_WHEAT
33	92.5	16.8	75	10RD_VIGON

## ALIGNMENTS

34	89.5	16.0	51	1	APF2_SINAL	P30232 sinapis alb
35	88	15.9	47	1	THG1_WHEAT	P20158 triticum ae
36	86.5	15.7	72	1	D230_PEA	O01783 pisum sativ
37	85.5	15.5	46	1	AX2_BETU	P82010 beta vulgar
38	85.5	15.5	51	1	APF1_SINAL	P30231 sinapis alb
39	84.5	15.3	46	1	AX1_BETU	P81493 beta vulgar
40	78	14.1	47	1	SIK1_SORBI	P21923 sorghum bic
41	77	13.9	70	1	DMYC_DROME	P41964 drosophila
42	76	13.8	47	1	PSD2_PEA	P81930 pisum sativ
43	76	13.8	47	1	THG2_MAIZE	P81009 zea mays (m
44	73.5	13.3	297	1	SGS4_DROME	O00725 drosophila
45	68	12.3	322	1	PSA_BRAE	Q9yhv4 brachydanto

RESULT 1  
THGF\_TOBAC  
ID THGF\_TOBAC  
AC P32026;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Flower-specific gamma-thionin precursor.  
GN FST.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RF SEQUENCE FROM A.A. STRAIN=CV, M88; TISSUE=Flower buds;  
RX MEDLINE=92357021; PubMed=1495489;  
RA Gu Q., Kawata E.E., Morse M.-J., Wu H.-M., Cheung A.Y.;  
RT "A flower-specific cDNA encoding a novel thionin in tobacco";  
RL Mol. Gen. Genet. 234:89-96(1992).  
CS -1- FUNCTION: INVOLVED IN FLORAL ORGANOGENESIS. MAY PLAY A PROTECTIVE  
ROLE IN FLOWERS BY PROTECTING THE REPRODUCTIVE ORGANS FROM  
POTENTIAL PATHOGEN ATTACK.  
CC -1- SUBCELLULAR LOCATION: POSSIBLY THE CELL WALL OR VACUOLE.  
CC -1- TISSUE SPECIFICITY: FLOWER. FOUND IN PETALS, STAMEN AND PISTILLS,  
BUT NOT IN SEEDS. IN PARTICULAR, ACCUMULATION IN A CONFIGURATION  
SURROUNDING THE INNER REPRODUCTIVE WHORLS.  
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES IN DEVELOPING FLOWERS AND ITS  
LEVEL DROPS AS FLOWERS MATURE.  
CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
DR EMBL: Z11748; CAA77806.1; -  
DR PIR: S23574; S23574.  
DR InterPro: IPR002318; Gamma-thionin.  
DR InterPro: IPR003614; Knott.  
DR Pfam: PF00304; Gamma-thionin; 1.  
DR ProDom: PD002594; G\_Purothionin; 1.  
DR SMART: SM00505; Knott; 1.  
DR PROSITE: PS00940; GAMMA\_THIONIN; 1.  
KW Plant defense; Plant toxin; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 105 FLOWER-SPECIFIC GAMMA-THIONIN.  
FT DISUPID 28 72 BY SIMILARITY.  
FT DISUPID 39 59 BY SIMILARITY.  
FT DISUPID 45 66 BY SIMILARITY.  
FT DISUPID 48 68 BY SIMILARITY.  
SQ SEQUENCE 105 AA; 11750 MW; DAFFSILF0631R03 CRCKA

Query Match 95.8%; Score 529; DB 1; Length 105;  
 Best Local Similarity 95.2%; Pred. No. 2.6e-49;  
 Matches 100; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 DB 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 QY 61 KIIRRCICTKPCVFDEKMTGTGAELIAEAKTIAALLBEEIMDN 105  
 DB 61 KIIRRCICTKPCVFDEKMTGTGAELIAEAKTIAALLBEEIMDN 105

## RESULT 2

THG1\_NICPA STANDARD; PRT; 106 AA.

AC 02415;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gamma-thionin 1 precursor.  
 GN THIO1.  
 OS Nicotiana paniculata.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_Taxid=62141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Komori T., Yamada S., Imaseki H.;  
 RT "A cDNA clone for gamma-thionin from Nicotiana paniculata."  
 RL (in) Plant Gene Register PGR97-132.  
 CC -1- INDUCTION: By salt stress.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB005250; BAA21325.1;  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knc01.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; G\_Purothionin; 1.  
 DR SMART; SM00505; Knc01; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; FALSE\_NEG.  
 KM Plant defense; Plant toxin; Signal.  
 FT SIGNAL  
 FT CHAIN  
 FT DISULFID 26 106  
 FT DISULFID 29 73  
 FT DISULFID 40 60  
 FT DISULFID 46 67  
 FT DISULFID 50 69  
 SQ SEQUENCE 106 AA; 11722 MW; 269850E116D775A4 CRC64;

Query Match 84.9%; Score 468.5; DB 1; Length 106;  
 Best Local Similarity 84.0%; Pred. No. 6.7e-43;  
 Matches 89; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 59  
 DB 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 QY 60 SKIIRRCICTKPCVFDEKMTGTGAELIAEAKTIAALLBEEIMDN 105  
 DB 61 SKIIRRCICTKPCVFDEKMTGTGAELIAEAKTIAALLBEEIMDN 106

Query Match 26.1%; Score 144; DB 1; Length 78;  
 Best Local Similarity 35.9%; Pred. No. 1.2e-08;  
 Matches 28; Conservative 14; Mismatches 30; Indels 6; Gaps 1;

QY 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 54  
 DB 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 QY 55 TDGHCKSLIRRCICTKPC 72  
 DB 61 IGCNCRAPRRRCCTRNC 78

## RESULT 4

THG4\_ARATH STANDARD; PRT; 73 AA.

AC 092018;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Gamma-thionin homolog PPT precursor.  
 GN Petunia integrifolia (Violet-flowered petunia) (Petunia inflata).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
 OC NCBI\_Taxid=4103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pistil;  
 RA MEDLINE=95036017; Pubmed=7948892;  
 RA Karunandaa B., Singh A., Rao T.H.;  
 RT "Characterization of a predominantly pistil-expressed gene encoding a  
 RT gamma-thionin-like protein of Petunia inflata."  
 RL Plant Mol. Biol. 26:459-464(1994).  
 CC -1- FUNCTION: May be involved in the defense of the pistil against  
 CC pathogen infection.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE PISTIL DURING  
 CC ALL STAGES OF FLOWER DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L27173; AAA64740.1;  
 DR PIR; S52634; S52634.  
 DR HSPR; P20307; IGRT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knc01.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; G\_Purothionin; 1.  
 DR SMART; SM00505; Knc01; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
 KM Plant defense; Signal.  
 FT SIGNAL  
 FT CHAIN  
 FT DISULFID 32 78  
 FT DISULFID 34 78  
 FT DISULFID 45 65  
 FT DISULFID 51 72  
 FT DISULFID 55 74  
 SQ SEQUENCE 78 AA; 8799 MW; 263DB3AC9D619265 CRC64;

Query Match 26.1%; Score 144; DB 1; Length 78;  
 Best Local Similarity 35.9%; Pred. No. 1.2e-08;  
 Matches 28; Conservative 14; Mismatches 30; Indels 6; Gaps 1;

QY 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 54  
 DB 1 MARSICFMAFALILAMLFVAYEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 QY 55 TDGHCKSLIRRCICTKPC 72  
 DB 61 IGCNCRAPRRRCCTRNC 78  
 RESULT 4  
 THG4\_ARATH STANDARD; PRT; 73 AA.  
 AC 092018;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

Thu Jan 29 11:57:49 2004

us-10-072-809b-17.rmpb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 09:22:12 / Search time 1169 Seconds  
(without alignments)  
1655.150 Million cell updates/sec

Title: US-10-072-809B-17  
Perfect score: 541  
Sequence: 1 atggctgcctcctctgtgctt.....tggttaaaaaaaaaaaaaa 541

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2356869 seqs, 1788235258 residues

Total number of hits satisfying chosen parameters: 4713738

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PTCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	541	100.0	541	US-10-072-809A-17	Sequence 17, Appl
2	240	44.4	223	US-10-072-809A-15	Sequence 15, Appl
3	223	41.2	240	US-10-072-809A-19	Sequence 19, Appl
4	216	39.9	216	US-10-072-809A-13	Sequence 13, Appl
5	189.8	35.1	564	US-09-777-347-1	Sequence 1, Appl
6	147.4	27.2	3528	US-09-777-347-2	Sequence 2, Appl
7	141	26.1	141	US-10-072-809A-7	Sequence 7, Appl
8	99	18.3	99	US-10-072-809A-11	Sequence 11, Appl
9	75	13.9	75	US-10-072-809A-9	Sequence 9, Appl
10	48	8.9	3673778	US-10-312-841-1	Sequence 1, Appl
11	46.2	8.5	516	US-10-178-213-82	Sequence 82, Appl
12	44.6	8.2	508	US-10-178-213-85	Sequence 85, Appl
13	44.6	8.2	578	US-10-178-213-292	Sequence 292, Appl
14	44.6	8.2	13326	US-10-311-455-1685	Sequence 1685, Ap
15	43.4	8.0	283	US-09-294-093B-4663	Sequence 4663, Ap

16	42.4	7.8	6042	US-10-311-455-1565	Sequence 1565, Ap
17	42.2	7.8	297	US-09-294-093B-4741	Sequence 4741, Ap
18	42.2	7.8	559	US-10-178-213-1	Sequence 1, Appl
19	42	7.8	459	US-10-178-213-322	Sequence 322, App
20	42	7.8	1692	US-10-349-680-155	Sequence 155, App
21	42	7.8	640681	US-09-790-988-1	Sequence 1, Appl
22	41.8	7.7	524	US-10-178-213-4	Sequence 4, Appl
23	41.8	7.7	6137	US-10-240-485-4	Sequence 4, Appl
24	41.6	7.7	1438	US-09-938-842A-4535	Sequence 4535, Ap
25	41.6	7.7	1438	US-09-938-842A-4535	Sequence 4535, Ap
26	41.2	7.6	468	US-10-178-213-316	Sequence 316, App
27	41.2	7.6	822900	US-10-252-798-1393	Sequence 277, App
28	41	7.6	640	US-10-178-213-55	Sequence 55, Appl
29	41	7.6	6665	US-10-240-485-3	Sequence 3, Appl
30	41	7.6	6665	US-10-240-485-3	Sequence 3, Appl
31	41	7.6	6665	US-10-239-676-3	Sequence 351, App
32	41	7.6	10433	US-10-311-455-351	Sequence 5390, Ap
33	40.8	7.5	608	US-09-814-353-5190	Sequence 11478, A
34	40.8	7.5	608	US-09-814-353-11478	Sequence 1809, App
35	40.8	7.5	6089	US-10-311-455-1809	Sequence 799, App
36	40.6	7.5	19659	US-10-311-455-739	Sequence 2, Appl
37	40.6	7.5	335913	US-09-754-853A-2	Sequence 3, Appl
38	40.6	7.5	335913	US-09-754-853A-3	Sequence 2, Appl
39	40.4	7.5	3673778	US-10-312-841-2	Sequence 17862, A
40	40.2	7.4	539	US-09-814-353-17862	Sequence 248, App
41	40.2	7.4	6074	US-10-240-453-248	Sequence 863, App
42	40.2	7.4	6131	US-10-311-455-863	Sequence 415, App
43	40	7.4	4190	US-10-178-213-424	Sequence 424, App
44	39.8	7.4	550	US-10-178-213-424	Sequence 2399, Ap
45	39.8	7.4	9219	US-10-311-455-2399	

#### ALIGNMENTS

US-10-072-809A-17  
Sequence 17, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072, 809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USN 60/267, 271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 17  
LENGTH: 541  
TYPE: DNA  
ORGANISM: Nicotiana glauca  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(318)  
US-10-072-809A-17

Query Match 100.0%; Score 541; DB 13; Length 541;  
Best Local Similarity 100.0%; Pred No. 9.6e-126; Indels 0; Gaps 0;  
Matches 541; Conservative 0; Mismatches 0

QY 1 ATGGCTGCTCTTGTGCTTCATAGGCAATTCATCTTGGCAAGATGCTTTGTGGC 60  
Db 1 ATGGCTGCTCTTGTGCTTCATAGGCAATTCATCTTGGCAAGATGCTTTGTGGC 60

QY 61 TATGAGGTGCACTGAGATGCAAAAGCAAGCAAAACATTCCTGGATATGCAAT 120  
Db 61 TATGAGGTGCACTGAGATGCAAAAGCAAGCAAAACATTCCTGGATATGCAAT 120

QY 121 ACCAAACCAACATGCAAGAAAGCTTATCAGTGAAGAAATTAAGTATGATGTAC 180  
Db 121 ACCAAACCAACATGCAAGAAAGCTTATCAGTGAAGAAATTAAGTATGATGTAC 180

122	ACCAAAACCAACGACGAAAAGCTTGATCAGTAGAAAATTACTGATGCTATTGAC	180	
QY	181	AAAATCCTAGAAAGTGCTTATGTACTAAGCATGTGTGATGGAAGATGACTAA	240
Db	181	AAAATCCTAGAAAGGCGCTATGTACTAAGCCATGTGTGTGATGGAAGATGACTAA	240
QY	241	ACAGAGCTGAATTTTGCTGAGAGAGCAAAAATTGGCTGAGCTTTGCTTGAAGA	300
Db	241	ACAGAGCTGGAATTTTGCTGAGAGAGCAAAAATTGGCTGAGCTTTGCTTGAAGA	300
QY	301	GAGATATGATPACTAATTAGATATGAGAAATTTAAGATGCAATATACATAT	360
Db	301	GAGATATGATPACTAATTAGATATGAGAAATTTAAGATGCAATATACATAT	360
QY	361	AAAGTTCTACCTTTCTAAAGTGAGCTATGTGTGTTTAAATGGCTTTAGTAC	420
Db	361	AAAGTTCTACCTTTCTAAAGTGAGCTATGTGTGTTTAAATGGCTTTAGTAC	420
QY	421	CTTTATATACCTTTAATATAGTGGCACTTCAATCTTTGTCATCTTGCACTA	480
Db	421	CTTTATATACCTTTAATATAGTGGCACTTCAATCTTTGTCATCTTGCACTA	480
QY	481	TTATTTGTGACTTTAATGAAAGAAGCTTGTATGTGCTTGGTTAAAAA	540
Db	481	TTATTTGTGACTTTAATGAAAGAAGCTTGTATGTGCTTGGTTAAAAA	540
QY	541	A 541	
Db	541	A 541	

```

RESULT 2
US-10-072-809A-15
: Sequence 15, Application US/10072809A
: Publication No. US20030217382A1
:
: GENERAL INFORMATION:
: APPLICANT: Anderson, Marilyn A., Lay, Fung T., Heath, Robyn, L.
: TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and v

```

Query Match	44.4%	Score 240;	DB 13;	Length 240;
Best Local Similarity	100.0%	Pred. NO. 2.2e-50;		
Matches 240; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

Db 181 TTGGCTGAGGAAAGCAAAAATTGGCTGACGTTGCTTGAAGAAGATATGATTAAC 240

RESULT 3  
US-10-072-809A-19  
; Sequence 19, Application US/10072809A  
; Publication No. US20030217382A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
; TITLE OF INVENTION: therefor  
; FILE REFERENCE: 18-01  
; CURRENT APPLICATION NUMBER: US/10/072, 809A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: USSN 60/267, 271  
; PRIOR FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 223  
; TYPE: DNA  
; ORGANISM: Nicotiana glauca  
US-10-072-809A-19

Query Match	41.2%;	Score 223;	DB 13;	Length 223;
Best Local Similarity	100.0%;	Pred. No. 3.9e-46;		
Matches 223;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 4
US-10-072-809A-13
; Sequence 13, Application US/10072809A
; Publication NO. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 216
; TYPE: DNA
; ORGANISM: Nicotiana glauca
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(216)
US-10-072-809A-13

```

Query Match	39.9%	Score 216	DB 13	Length 216
Best Local Similarity	100.0%	Pred. No. 2.2e-44		
Matches 216; Conservative	0	Mismatches 0	Indels 0	Gaps 0

QY 1 ATGGCTGCTCTTGTGCTTATGCAATTTGCTATCTTGGAAGAGTCTTTGTTGCC 60  
DB 1 ATGGCTGCTCTTGTGCTTATGCAATTTGCTATCTTGGAAGAGTCTTTGTTGCC 60  
QY 61 TATGAGTGCAGCTAGAGAAATGCAAAACAGAAACACATTTCTTGAAATATGACT 120  
DB 61 TATGAGTGCAGCTAGAGAAATGCAAAACAGAAACACATTTCTTGAAATATGACT 120  
QY 121 ACCAACAACACATGAGAAAGCTGTATCACTGAGAAATTTACTGATGTCATGTAGC 180  
DB 121 ACCAACAACACATGAGAAAGCTGTATCACTGAGAAATTTACTGATGTCATGTAGC 180  
QY 181 AAAATCTCAGAAAGTGCCTTATGTAATAAGCATGT 216  
DB 181 AAAATCTCAGAAAGTGCCTTATGTAATAAGCATGT 216

## RESULT 5

US-09-777-347-1  
Sequence 1, Application US/09777347  
Patent No. US20010014977A1  
GENERAL INFORMATION:  
APPLICANT: McBride, Kevin E.  
Stalker, David M.  
TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calgene, Inc.  
STREET: 1920 Fifth Street  
CITY: Davis  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1 (a)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/777,347  
FILING DATE: 05-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,087  
FILING DATE: <Unknown>  
APPLICATION NUMBER: USN 07/554,195  
FILING DATE: 17-JUL-90  
APPLICATION NUMBER: USN 07/382,518  
FILING DATE: 19-JUL-89  
ATTORNEY/AGENT INFORMATION:  
NAME: Carl J. Schwedler  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: CGNE 91-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (916) 753-6313  
TELEFAX: (916) 753-1510  
INFORMATION FOR SEQ ID NO: 1  
SEQUENCE DESCRIPTION: SEQ ID NO: 1  
US-09-777-347-1

Query Match 35.1%; Score 189.8; DB 9; Length 564;  
Best Local Similarity 68.9%; Pred. No. 1,3e-37;  
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;  
QY 1 ATGGCTGCTCTTGTGCTTATGCAATTTGCTATCTTGGAAGAGTCTTTGTTGCC 60  
DB 40 ATGGCTGCTCTTGTGCTTATGCAATTTGCTATCTTGGAAGAGTCTTTGTTGCC 99

QY 61 TATGAGTGCAGCTAGAGAA--TGCAAAACAGAAAGCAACATTTCTTGAAATATGC 117  
DB 100 TATGAGTGCAGCTAGAGAAATTTGCAAAAGCAAGCAAACTTTCCAGATATATGT 159  
QY 118 ATTACCAACCAACATGAGAAAGCTGTATCACTGAGAAATTTACTGATGTCATGT 177  
DB 160 TTTATGACTCATCATGTAAGAAATATGTAATCAAGAAATTTACTGATGTCATGT 219  
QY 178 AGCAAAATCTCAGAAAGTGCCTATGTAATAAGCATGTGTGTGATGAGAAATGACT 237  
DB 220 AGCAAACTCCAAAGAAAGTGTATGCACTAAGCCATGTGTAT--TGCAAAATCTCA 276  
QY 238 AAAACGAGCTGAATTTTGGCTGAGAGCAAAACCTTGGCTGCACCTTGTGTTGA 297  
DB 277 AGTGAAGTTAAACCACTTGGGTGAGAGCAAAACCTTAAGTGAAGTGTGCTTGA 336  
QY 298 GAAGATATATGATTAATAATTAGATTAGAGAAATTAAGAT-----GCA 346  
DB 337 GAAGATATATGATTAATAATTAGATTAGAGTAAATTAAGATTTGAGTCAAA 396  
QY 347 GTATCAACATTAATAAGTTTCTTACTTCTTAAAGTATGCT--TATGTTGCTTT 403  
DB 397 AAAACAAATTAATAAGTGTGCTTCTTATTAAGGTAAGTGTGATGTTGTTAG 456  
QY 404 AATTGCTTTTATGACCTTTTATTAACCTTAAATAAGTGTGACATTCAT 456  
DB 457 TATGGCTTATGATGACATTTGACATTAATAAGTGTGACATTCAT 509

## RESULT 6

US-09-777-347-2  
Sequence 2, Application US/09777347  
Patent No. US20010014977A1  
GENERAL INFORMATION:  
APPLICANT: McBride, Kevin E.  
Stalker, David M.  
TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calgene, Inc.  
STREET: 1920 Fifth Street  
CITY: Davis  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1 (a)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/777,347  
FILING DATE: 05-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,087  
FILING DATE: <Unknown>  
APPLICATION NUMBER: USN 07/554,195  
FILING DATE: 17-JUL-90  
APPLICATION NUMBER: USN 07/382,518  
FILING DATE: 19-JUL-89  
ATTORNEY/AGENT INFORMATION:  
NAME: Carl J. Schwedler  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: CGNE 91-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (916) 753-6313  
TELEFAX: (916) 753-1510  
INFORMATION FOR SEQ ID NO: 2  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3528 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2  
US-09-777-347-2

Query Match 27.2%; Score 147.4; DB 9; Length 3528;  
Best Local Similarity 67.1%; Pred. No. 1.5e-26;  
Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

QY 82 TGCAGAAACAGAAAGCAACACATTTCCGGAATATGATTAACAACCCAGCATGAGAAA 141  
DB 2942 TGCAGAGACCAACAGCAACATTTCCAGGATATGTTATGAGCTCATGAGAAA 3001  
QY 142 GCTTGATCATGAGAAATTTACTGATGTCATGTCGCAAAATCCCAAGAGTGGCTA 201  
DB 3002 TATTGTATCAAGAGAAATTTACTGTCGAGCATGTGCAAACTCCAAAGAGTGTCTA 3061  
QY 202 TGTACTAAGCCATGTGTGTTGATGAGAGATGACTAAAACAGAGCTGAAATTTGGCT 261  
DB 3062 TGCACATAGCCATGTGTATTT--TGACAAATCTCAAGTAAAGTAAAGCACTTGGGT 3118  
QY 262 GAGGAGCAAAACCTTTGCTGCGCTTTGCTTGAAGAAGATTAAGTAACTAATTA 321  
DB 3119 GAGGAGCAAAACCTTAAAGTGAAGTGTGCTTGAAGAAGATTAAGTGAAGTAAATTA 3178  
QY 322 GAGATTAAGAGAAATTAAGAT-----GCAGTATCACACATATTAAGTTCTA 370  
DB 3179 TTAAGTGAAGTAAATAGATTTTGAAGTCAAAAACAAATTAATTAAGTGTGC 3238  
QY 371 CTTTCTTAAAGAGTACG--TAATGTTGTTTAAATGCTTTAGTACCTTTAT 427  
DB 3239 CTTTCTTAAAGAGTACGTTGATGTTGTTAGTAAATGAGCTTAAATGATGGA 3298  
QY 428 TAACTTTAAATTAAGTGTGCACTTCAT 456  
DB 3299 CACATTAATTAAGTGTGCACTTCAT 3327

RESULT 7  
US-10-072-809A-7  
Sequence 7, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 7  
LENGTH: 141  
TYPE: DNA  
ORGANISM: Nicotiana glauca  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(141)  
US-10-072-809A-7

Query Match 26.1%; Score 141; DB 13; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.2e-25;  
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 AGAAGATGCAAAACAGAAAGCAACATTTCTGAAATATGATTAACAACCCAGCATGC 135  
DB 1 AAGAGATGCAAAACAGAAAGCAACATTTCTGAAATATGATTAACAACCCAGCATGC 60  
QY 136 AAGAAAGCTTGATGATGAGAAATTTACTGATGCTATGAGAAATCTCTGAGAG 195

DB 61 AAGAAAGCTTGATGATGAGAAATTTACTATGATGATGATGAGAAATCTCTGAGAG 120  
QY 196 TGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 216  
DB 121 TGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 141

RESULT 8  
US-10-072-809A-11  
Sequence 11, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 99  
TYPE: DNA  
ORGANISM: Nicotiana glauca  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(99)  
US-10-072-809A-11

Query Match 18.3%; Score 98; DB 13; Length 99;  
Best Local Similarity 100.0%; Pred. No. 3.5e-15;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 GTGTTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276  
DB 1 GTGTTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
QY 277 TTGGCTGACGCTTTGCTTGAAGAGATTAATGATTAAC 315  
DB 61 TTGGCTGACGCTTTGCTTGAAGAGATTAATGATTAAC 99

RESULT 9  
US-10-072-809A-9  
Sequence 9, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 75  
TYPE: DNA  
ORGANISM: Nicotiana glauca  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(75)  
US-10-072-809A-9

Query Match 13.9%; Score 75; DB 13; Length 75;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGCTCTTGTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60



Thu Jan 29 11:57:49 2004

us-10-072-809b-17.rnpb

Page 5

Db 1 ATGCGCTGCGCTCTGTGTGCTTCATGGCAATTGCTCATCTTGGCAAGATGCTCTTTGTGCC 63

Qy 61 TATAGGTGCAGCT 75

Db 61 TATAGGTGCAGCT 75

```

RESULT 10
US-10-312-841-1
: Sequence 1, Application US/10312841
: Publication No. US20030186277A1
: GENERAL INFORMATION:
: APPLICANT: Epigenomics AG
: TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
: FILE REFERENCE: E01/1208/WO
: CURRENT APPLICATION NUMBER: US/10/312,841
: CURRENT FILING DATE: 2002-12-30
: NUMBER OF SEQ ID NOS: 2
: SEQ ID NO 1
: LENGTH: 3673778
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
: NAME/KEY: unsure
: LOCATION: (3294164)
: US-10-312-841-1

```

	Query Match	8.9%	Score	48;	DB	13;	Length	3673778;		
	Best Local Similarity	50.9%;	Pred.	No. 4;						
	Matches	114;	Conservative	0;	Mismatches	110;	Indels	0;	Gaps	0;
Qy	299	AAGAGATATGCACTAATTAAGACATTGAAGAATTTAGAGTGCAGATCACACATA	358							
Dd	66836	AAAGTATTAAGAAAAATTAATTAAGAAAAGATATTGATTTAAATTTGGAAATA	66899							
Qy	359	ATAAGTTTCACCTTCTTAAAGGTGACTGATGTTTGTTAAATGCGCTTTAGTA	418							
Dd	66896	ATAATTTAAATTTGTTTTTAATAGATGATGAATTTTGGATATTTTTTTTTTAAAG	66955							
Qy	419	GCCTTTATTAACCTTTAAATAGGTGGCACCITTCACCTTTGTGCCATCTTGACCTAA	478							
Dd	66956	GTTGATTTAATTTTTTTTTTTGGTAGTTGGGTTTAAAGATGATTTTTGATATAG	67015							
Qy	479	GTTTATTTGACCTTTTAATGAATGACCTTCATGTCCTTT	522							
Dd	67016	GTTTATTTTGTGTTTATTTAATTAATTTTGAATTTGAAT	67059							

RESULT 11  
 US-10-178-213-82  
 : Sequence 82, Application US/10178213  
 : Publication No. US20030041348A1  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Simmons, Carl R.  
 : APPLICANT: Navarrio Acevedo, Pedro A.  
 : APPLICANT: Harvell, Leslie  
 : APPLICANT: Cahoon, Rebecca  
 : APPLICANT: McCutchen, Billy Fred  
 : APPLICANT: Lu, Albert  
 : APPLICANT: Hermann, Rafael  
 : APPLICANT: Wong, James  
 :  
 : TITLE OF INVENTION: Defensein Polynucleotides and Methods of  
 : TITLE OF INVENTION: Use  
 : FILE REFERENCE: 3578/246703  
 :  
 : CURRENT FILING DATE: 2002-06-21  
 :  
 : PRIOR APPLICATION NUMBER: 60/300,152  
 : PRIOR FILING DATE: 2001-06-22  
 :  
 : PRIOR APPLICATION NUMBER: 60/300,241  
 : PRIOR FILING DATE: 2001-06-22

```

? NUMBER OF SEQ ID NOS: 469
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 82
? LENGTH: 516
? TYPE: DNA
? ORGANISM: Beta vulgaris
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (31)...(267)
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: (121)...(261)
US-10-178-213-82

```

	Query Match	8.5%	Score 46.2	DB 15	Length 516
	Best Local Similarity	57.1%	Pred. No. 0.14		
	Matches	84	Conservative	0	Mismatches 63
					Indels 0
					Gaps 0
QY	71	AAAGCTAGAGATGACAAAACAGAAAGCAACACATTTCTGTGAAATATGCAATTACCAACAC	130		
DB	116	AAACAGAGACATGTATGACTCCAGTACACGATTCAAGGGAAATATGGTTAGTAGAGAA	175		
QY	131	CATGCAGAAAGCTTGTCATCAGTGGAGAAATTACTATGTGTCATTGTAGCAAAATCTCTA	190		
DB	176	ATTGTAAATCTCTTGCCACACTGAGAGATTTCTGTGGAGAACGTGTCAAGCTTTCGTA	235		
QY	191	GAAGTGCCATGTATCTAAGCATGTG	217		
DB	236	GAAGATGATGTGCACTAAGCCTTGCG	262		

```

RESULT 12
US-10-178-213-85
; Sequence 85, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensein Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Beta vulgaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)...(267)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (121)...(261)
US-10-178-213-85

```

Query Match 8.2%; Score 44.6; DB 15; Length 508;  
 Best Local Similarity 56.5%; Pred. No. 0.33;  
 Match 83; Conservative 0; Mismatches 64; Indels 0; Gaps 0

71 AAGCTGAAATGCAAAACAGAAAGCAACATTCTCGAATATGATTCATCCAAACGAC 130

Db 116 AAGCAGAGCATGTATGACTCCAGTCACCGAGGGAATATCGTTAGTAGTGA 175  
Cy 131 CATGCAAGAAAGCTTGATGATGAGAAATTTACTGATGCTTGTAGCAAAATCTCA 190  
Db 176 ATGTGATCTGCTTGCCACACTGAGAGATTTCCTGGAGAGAGTGTCAAGGCTTCGTA 235  
Cy 191 GAAGGTGCTTATGATCTTACCAAGCATGTG 217  
Db 236 GAAGATGATGTGACACTGAGGCTTGGC 262

## RESULT 13

US-10-178-213-292  
Sequence 292, Application US/10178213  
Publication No. US20030041348A1  
GENERAL INFORMATION:  
APPLICANT: Simmons, Carl R.  
APPLICANT: Navarro Acevedo, Pedro A.  
APPLICANT: Harvell, Leslie  
APPLICANT: Cahoon, Rebecca  
APPLICANT: McCutchen, Billy Fred  
APPLICANT: Lu, Albert  
APPLICANT: Herimann, Rafael  
TITLE OF INVENTION: Defensin Polynucleotides and Methods of  
TITLE OF INVENTION: Use  
FILE REFERENCE: 35718/246703  
CURRENT APPLICATION NUMBER: US/10/178, 213  
CURRENT FILING DATE: 2002-06-21  
PRIOR APPLICATION NUMBER: 60/300,152  
PRIOR FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/300,241  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 469  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 292  
LENGTH: 578  
TYPE: DNA  
ORGANISM: Beta vulgaris  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (62)...(301)  
FEATURE:  
NAME/KEY: mat.peptide  
LOCATION: (152)...(292)  
US-10-178-213-292

Query Match 8.2%; Score 44.6; DB 15; Length 578;  
Best Local Similarity 56.5%; Pred. No. 0.37;  
Matches 83; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Cy 71 AAGCTAGAGATGCAAAAGCAAGCAACACATTTCTCGAATATGATCTTACCAACAC 130  
Db 147 AAGCAGAGACTGTGAACTCCAGCCAGAGGTGAGGAGTATGTTGAAAGAGAA 206  
Cy 131 CATGCAAGAAAGCTTGATGATGAGAAATTTACTGATGCTTGTAGCAAAATCTCA 190  
Db 207 ATGTGATCTGCTTGCCACACTGAGAGATTTCCTGGAGAGAGTGTCAAGGCTTCGTA 266  
Cy 191 GAAGGTGCTTATGATCTTACCAAGCATGTG 217  
Db 267 GAAGATGATGTGACACTGAGGCTTGGC 262

## RESULT 14

US-10-311-455-1685  
Sequence 1685, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BRILIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect

TITLE OF INVENTION: cytosine methylation  
FILE REFERENCE: 5013, 1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 1685  
LENGTH: 13326  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1685

Query Match 8.2%; Score 44.6; DB 13; Length 13326;  
Best Local Similarity 47.3%; Pred. No. 1.8;  
Matches 134; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Cy 254 TTTGGCTGAGAGCAAAAGCTTGCTGAGCTTGTGCTGAAGAGATTAATGATA 313  
Db 10474 TTAATTTTAAAGATTAATTAATTTGTTGTTAAGTTGTTGTTGTTAATTTGTT 10533  
Cy 314 ACTAATAGAGATTAGAAATTAAGATGACATGATACACATTAATTAAGTTCTACT 373  
Db 10534 ACGTAGTAATAGAAATTAATTAAGATTAATTAAGATTAATTAATTAAGTAGT 10593  
Cy 374 TTTCTAAAGTACATGATGTTGTTTAAATGCTTTAGTACCTTTTAACTACT 433  
Db 10594 ATGTAGTAGTAATAGATGATTTAGTTTGAATTTGTTGTTGTTTAAATAGGA 10653  
Cy 434 TTAATAGTGTGAGCACTTCAATCTTGTGCAATCTTGCACTAAGTTTATTTGTACT 493  
Db 10654 TTGTTTGTGTTTAAAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGT 10713  
Cy 494 TTTAATGAAATGACCTTCTAATGCTTGTGTTAAATTAATTAATTAATTAATTA 536  
Db 10714 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 10756

## RESULT 15

US-09-294-093B-4663  
Sequence 4663, Application US/09294093B  
Patent No. US20010051335A1  
GENERAL INFORMATION:  
APPLICANT: Ito, Laura, Y.  
APPLICANT: Sherman, Bradley, K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
FILE REFERENCE: PL-0009 US  
CURRENT APPLICATION NUMBER: US/09/294,093B  
CURRENT FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: 60/082,567  
PRIOR FILING DATE: April 21, 1998  
NUMBER OF SEQ ID NOS: 6207  
SOFTWARE: PERL Program  
SEQ ID NO 4663  
LENGTH: 283  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20010051335A1 700354759H1  
NAME/KEY: unsure  
LOCATION: 14, 100  
OTHER INFORMATION: a, t, c, g, or other  
US-09-294-093B-4663

Query Match 8.0%; Score 43.4; DB 9; Length 283;

QY 5 CCTATAGGTGCAGCTATAGAAATCAAAACGAAAGCAGACATTTCTGSAATATGCA 118  
Db 111 CCATATTTGGCAGCAGGAAAGGAAATGGAGAACTACATGAGACCGATTGTTGGGGCATGCA 170  
QY 119 TTACCAACACCCATCATCAGAAAAGCTTGATCAGTGAATATTTACTGATGGTCATGTGA 178  
Db 171 TGATCGAGACACTGTGGCCATATGTGTGGCCGGGTGAAGGCGTTCTTGGGCGGAGATGCA 230  
QY 179 GCAAAATCCTCAGAAAGTGCTTATGTACTTAAAGCATG 215  
Db 231 GCAAGCTTCCGCGCGCTGATCTGTGACTAGCAAGTG 267

Search completed: January 28, 2004, 11:23:53  
Job time : 1177 secs

Thu Jan 29 11:57:49 2004

us-10-072-809b-17.rn1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 08:44:46 ; Search time 56 Seconds  
(without alignments)

4264.076 Million cell updates/sec

Title: US-10-072-809B-17

Perfect score: 541  
Sequence: 1 atggctcgctcttgcctt.....tggtaaaaaaaaaaaaaa 541

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfil1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189.8	35.1	564	3	US-08-984-320-1
2	189.8	35.1	564	3	US-08-487-087A-1
3	189.8	35.1	564	6	US-08-397-653B-1
4	189.8	35.1	564	6	5175095-1
5	147.4	27.2	3528	3	US-08-984-320-2
6	147.4	27.2	3528	3	US-08-487-087A-2
7	147.4	27.2	4383	4	US-08-397-653B-2
8	147.4	27.2	4383	6	5175095-4
9	147.4	27.2	4383	6	5177307-1
10	95.6	17.7	506	4	US-09-442-631-1
11	45.4	8.4	7218	1	US-08-232-463-14
12	42	7.0	53332	4	US-09-601-198-63
13	38	6.9	1738	1	US-08-472-028A-3
14	37.2	6.9	1738	2	US-08-808-931-3
15	37.2	6.9	1738	3	US-08-808-931-3
16	37.2	6.9	1738	3	US-08-808-931-3
17	37.2	6.9	1738	3	US-09-050-603A-3
18	37.2	6.9	1738	3	US-09-102-420B-3
19	37.2	6.9	1738	3	US-09-071-1296-3
20	37.2	6.9	1738	3	US-09-196-268-3
21	37.2	6.9	1738	3	US-09-015-683-3
22	37.2	6.9	1738	4	US-09-191-998-3
23	37.2	6.9	1738	4	US-09-497-698-3
24	37.2	6.9	19124	2	US-08-487-826B-13
25	36.6	6.8	225	4	US-09-442-631-3
26	36.6	6.8	10607	1	US-08-078-090-3
27	36.2	6.7	3000	2	US-08-680-395-1

C 28	36	6.7	1939	1	US-07-715-751B-2	Sequence 2, Appli
C 29	35.4	6.5	19124	2	US-08-487-826B-13	Sequence 13, Appli
C 30	35	6.5	6701	4	US-09-611-659A-1	Sequence 1, Appli
C 31	35	6.5	580073	4	US-08-545-528D-1	Sequence 1, Appli
C 32	35	6.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 33	35	6.5	1664976	4	US-08-916-421B-1	Sequence 2, Appli
C 34	34.8	6.4	2943	4	US-09-503-922-2	Sequence 75, Appli
C 35	34.6	6.4	2028	2	US-08-933-750C-75	Sequence 1, Appli
C 36	34.6	6.4	2028	3	US-09-234-613-75	Sequence 1, Appli
C 37	34.6	6.4	3000	1	US-08-184-252A-1	Sequence 1, Appli
C 38	34.6	6.4	3000	5	PCT-US95-00601-1	Sequence 1, Appli
C 39	34.6	6.4	162450	4	US-09-345-882-1	Sequence 19, Appli
C 40	34.4	6.4	1786	1	US-07-920-430-19	Sequence 9, Appli
C 41	34.4	6.4	1786	1	US-08-066-299-9	Sequence 1, Appli
C 42	34.4	6.4	1786	1	US-08-265-047-1	Sequence 19, Appli
C 43	34.4	6.4	1786	1	US-08-251-464-19	Sequence 20, Appli
C 44	34.4	6.4	1786	4	US-08-926-522-20	Sequence 19, Appli
C 45	34.4	6.4	1786	5	PCT-US92-01364-19	

#### ALIGNMENTS

RESULT 1  
US-08-984-320-1  
Sequence 1, Application US/08984320  
Patent No. 622097  
GENERAL INFORMATION:  
APPLICANT: McBridge, Kevin E.  
APPLICANT: Stalker, David M.  
TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calgene, Inc.  
STREET: 1920 Filth Street  
CITY: Davis  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1 (a)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/984,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,087  
FILING DATE: 07-JUN-95  
APPLICATION NUMBER: USN 07/998,158  
FILING DATE: 29-DEC-92  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USN 07/554,195  
FILING DATE: 17-JUL-90  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USN 07/382,518  
FILING DATE: 19-JUL-89  
ATTORNEY/AGENT INFORMATION:  
NAME: Donna E. Scherer  
REGISTRATION NUMBER: 34,719  
NAME: Carl V. Schwedler  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: CGNE 91-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (916) 753-6313  
TELEFAX: (916) 753-1510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 564 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double



ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.0  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/397,653B  
FILING DATE: 28-FEB-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth Lassen  
REGISTRATION NUMBER: 31,845  
NAME: Donna E. Scherer  
REGISTRATION NUMBER: 34,719  
NAME: Carl J. Schwedler  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: CGNE 112  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (916) 753-6313  
TELEFAX: (916) 753-1510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 564 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
US-08-397-653B-1

Query Match 35.1%; Score 189.8; DB 4; Length 564;  
Best Local Similarity 68.9%; Pred. No. 7.9e-42;  
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 ATGGCTGCTCTTGGCTTCATGCGATTTGCTATCTTGGCAAGATGCTCTTTGTTGCC 60  
DB 40 ATGGCTGCTCTTGGCTTCATGCGATTTTGGCTTGGCAATGATGCTCTTTGTTACC 99  
QY 61 TATGAGGTGCAAGCTAGAGAA--TGCAAAACAGAAAGACACATTTCTCGAATATGC 117  
DB 100 TATGAGGTGCAAGCTAGAGAAATTTGCAAAAGCACCAACCACTTCCAGGATTATGT 159  
QY 118 ATTACCAACCAACATGACGAAAGCTGTATCAGTGAAGAAATTACTGATGTCATGT 177  
DB 160 TTTATGACTCATCATGTGAGAAAATTTGATCAAGAGAAATTTACTGGTGACATGT 219  
QY 178 AGCAAAATCTCTGAGAGGTGCTATGTACTTAAGCAATGTGTGTGATGAGAAATGACT 237  
DB 220 AGCAAACTCCAAAGAGGTGCTATGACCTAAGCCATGTGTAT--TGACAAATCTCA 276  
QY 238 AAAACAGAGCTGAATTTTGGCTGAGAGAGCAAAACCTTGGCGAGCTTGGTTGAA 297  
DB 277 AGTGAAGTTAAAGCACTTTGGGTGAGAGCAAAACCTCTAGTGAAGTTGTGCTTAA 336  
QY 298 GAAGAGATATGATTAATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346  
DB 337 GAAGAGATATGATGAGATTAATTAATTAAGATTAGAGAAATTTGAGTGTCAAAA 396  
QY 347 GTATCACAATAATAAGTTTCTACTCTTTCTTAAGAGTAC--TAATGTGTGTTT 403  
DB 397 AAAACAAATTAATAAAGTTGCTTTCTTAATTAAGAGTGTGAGTGTGTGAG 456  
QY 404 AATTGGCTTTAGAGCTTTTATTAACACTTAAATTAAGTGTGAGACTTCAT 456  
DB 457 TATGGCTTATGATGACCATTTGACACATTAATAAGTTGTGACACATCAT 509

RESULT 4  
5175095-1  
; Patent No. 5175095  
; APPLICANT: Martineau, Belinda M.; Houck, Catherine M.  
; TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS  
; NUMBER OF SEQUENCES: 9

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/554,195  
FILING DATE: 17-JUL-1990  
SEQ ID NO: 1  
LENGTH: 564  
5175095-1

Query Match 35.1%; Score 189.8; DB 6; Length 564;  
Best Local Similarity 68.9%; Pred. No. 7.9e-42;  
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 ATGGCTGCTCTTGGCTTCATGCGATTTGCTATCTTGGCAAGATGCTCTTTGTTGCC 60  
DB 40 ATGGCTGCTCTTGGCTTCATGCGATTTTGGCTTGGCAATGATGCTCTTTGTTACC 99  
QY 61 TATGAGGTGCAAGCTAGAGAA--TGCAAAACAGAAAGACACATTTCTCGAATATGC 117  
DB 100 TATGAGGTGCAAGCTAGAGAAATTTGCAAAAGCACCAACCACTTCCAGGATTATGT 159  
QY 118 ATTACCAACCAACATGACGAAAGCTGTATCAGTGAAGAAATTACTGATGTCATGT 177  
DB 160 TTTATGACTCATCATGTGAGAAAATTTGATCAAGAGAAATTTACTGTGACATGT 219  
QY 178 AGCAAAATCTCTGAGAGGTGCTATGTACTTAAGCAATGTGTGTGATGAGAAATGACT 237  
DB 220 AGCAAACTCCAAAGAGGTGCTATGACCTAAGCCATGTGTAT--TGACAAATCTCA 276  
QY 238 AAAACAGAGCTGAATTTTGGCTGAGAGCAAAACCTTGGCGAGCTTGGTTGAA 297  
DB 277 AGTGAAGTTAAAGCACTTTGGGTGAGAGCAAAACCTCTAAGTGAAGTTGTGCTTGA 336  
QY 298 GAAGAGATATGATTAATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346  
DB 337 GAAGAGATATGATGAGATTAATTAATTAAGATTAGAGAAATTTGAGTGTCAAAA 396  
QY 347 GTATCACAATAATAAGTTTCTACTCTTTCTTAAGAGTAC--TAATGTGTGTTT 403  
DB 397 AAAACAAATTAATAAAGTTGCTTTCTTAATTAAGAGTGTGAGTGTGTGAG 456  
QY 404 AATTGGCTTTAGAGCTTTTATTAACACTTAAATTAAGTGTGAGACTTCAT 456  
DB 457 TATGGCTTATGATGACCATTTGACACATTAATAAGTTGTGACACATCAT 509

RESULT 5  
US-08-984-320-2  
; Sequence 2, Application US/08984320  
; Patent No. 622097  
; GENERAL INFORMATION:  
; APPLICANT: McBride, Kevin E.  
; APPLICANT: Stalker, David M.  
; TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Calgene, Inc.  
; STREET: 1920 Fifth Street  
; CITY: Davis  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 95616

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1 (a)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/984,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,087  
FILING DATE: 07-JUN-95  
APPLICATION NUMBER: USSN 07/998,158



Patent No. 6329570  
 GENERAL INFORMATION:  
 APPLICANT: Martineau, Belinda  
 TITLE OF INVENTION: COTTON MODIFICATION USING  
 TITLE OF INVENTION: OVARY-TISSUE TRANSCRIPTIONAL  
 TITLE OF INVENTION: FACTORS  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Calgene, Inc.  
 STREET: 1920 Fifth Street  
 CITY: Davis  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 95616  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB  
 COMPUTER: Apple Macintosh  
 OPERATING SYSTEM: Macintosh 7.0  
 SOFTWARE: Microsoft Word 5.1a  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/397,653B  
 FILING DATE: 28-FEB-1995  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elizabeth Lassen  
 REGISTRATION NUMBER: 31,845  
 NAME: Donna E. Scherer  
 REGISTRATION NUMBER: 34,719  
 NAME: Carl J. Schwedler  
 REGISTRATION NUMBER: 36,924  
 REFERENCE/DOCKET NUMBER: CGNE 112  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (916) 753-6313  
 TELEFAX: (916) 753-1510  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4383 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 US-08-397-653B-2

Query Match 27.2%; Score 147.4; DB 4; Length 4383;  
 Best Local Similarity 67.1%; Pred. No. 3.3e-30;  
 Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

QY 82 TGCAGAAACGAAAGCAACATTTCTGGAATATGCAATACCAACCCATGCGAAAA 141  
 DB 2942 TGCAGAAACGAAAGCAACATTTCTGGAATATGCAATACCAACCCATGCGAAAA 3001  
 QY 142 GCTTGATCAGTGAAGAAATTTACTGATGTCATTTAGCAAAATCCTAGAAGTGCTTA 201  
 DB 3002 TATTGATCAAGAGAAATTTACTGATGTCATTTAGCAAAATCCTAGAAGTGCTTA 3061  
 QY 202 TGTACTAAGCCATGTGTGTTTATGAGAAAGTGAATCAAAAACGAGCTGAAATTTGGCT 261  
 DB 3062 TGCACATAGCCATGTGTATTT--TGACAAATCTCAAGTGAAGTTAAAGCAATTTGGGT 3118  
 QY 262 GAGGAAGCAAAACTTTGGCTGAGCTTGTGTAAGAAAGATATAGTAACTAATTA 321  
 DB 3119 GAGGAAGCAAAACTTTAGTGAAGTGTGCTTGAAGAAAGATATAGTGAAGTAA 3178  
 QY 322 GAGATTAGAAGAAATTAAGAT-----GCAGTATCACACATATAAAGTTCTA 370  
 DB 3179 TTAAGAGAGTTAAATTAAGATTTGAGTGTCAAAAAACAAATTAATAAGTGTC 3238  
 QY 371 CCTTTCTTAAAGTGAAGC---TAATGTGTGTTTAAATGCTTTTGTAGCCTTTAT 427  
 DB 3239 CTTTCTTTTAAAGGAGCTTGTGATGTTGTGTATGATTTGGCTATAGTACCATTTGA 3298  
 QY 428 TACACTTAAATTAAGTGTGCACTTCAT 456

DB 3299 CACATTAAATTAAGTTGTGACATCATCATT 3327  
 RESULT 8  
 Patent No. 5175095-4  
 APPLICANT: Martineau, Belinda M.; Houck, Catherine M.  
 TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS  
 NUMBER OF SEQUENCES: 9  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/554,195  
 FILING DATE: 17-JUL-1990  
 SEQ ID NO: 4  
 LENGTH: 4383

Query Match 27.2%; Score 147.4; DB 6; Length 4383;  
 Best Local Similarity 67.1%; Pred. No. 3.3e-30;  
 Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

QY 82 TGCAGAAACGAAAGCAACATTTCTGGAATATGCAATACCAACCCATGCGAAAA 141  
 DB 2942 TGCAGAAACGAAAGCAACATTTCTGGAATATGCAATACCAACCCATGCGAAAA 3001  
 QY 142 GCTTGATCAGTGAAGAAATTTACTGATGTCATTTAGCAAAATCCTAGAAGTGCTTA 201  
 DB 3002 TATTGATCAAGAGAAATTTACTGATGTCATTTAGCAAAATCCTAGAAGTGCTTA 3061  
 QY 202 TGTACTAAGCCATGTGTGTTTATGAGAAAGTGAATCAAAAACGAGCTGAAATTTGGCT 261  
 DB 3062 TGCACATAGCCATGTGTATTT--TGACAAATCTCAAGTGAAGTTAAAGCAATTTGGGT 3118  
 QY 262 GAGGAAGCAAAACTTTGGCTGAGCTTGTGTAAGAAAGATATAGTAACTAATTA 321  
 DB 3119 GAGGAAGCAAAACTTTAGTGAAGTGTGCTTGAAGAAAGATATAGTGAAGTAA 3178  
 QY 322 GAGATTAGAAGAAATTAAGAT-----GCAGTATCACACATATAAAGTTCTA 370  
 DB 3179 TTAAGAGAGTTAAATTAAGATTTGAGTGTCAAAAAACAAATTAATAAGTGTC 3238  
 QY 371 CCTTTCTTAAAGTGAAGC---TAATGTGTGTTTAAATGCTTTTGTAGCCTTTAT 427  
 DB 3239 CTTTCTTTTAAAGGAGCTTGTGATGTTGTGTATGATTTGGCTATAGTACCATTTGA 3298  
 QY 428 TACACTTAAATTAAGTGTGCACTTCAT 456  
 DB 3299 CACATTAAATTAAGTTGTGACATCATCATT 3327

RESULT 9  
 Patent No. 5177307-1  
 APPLICANT: HOUCK, CATHERINE M.; PEAR, JULIE R.; MARTINEAU, BELINDA M.; HATT, WILLIAM  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
 MODULATIONS OF ENDOGENOUS CYTOKININ LEVELS  
 NUMBER OF SEQUENCES: 6  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/554,196  
 FILING DATE: 17-JUL-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 382,802  
 FILING DATE: 19-JUL-1989  
 APPLICATION NUMBER: 188,361  
 FILING DATE: 29-APR-1988  
 APPLICATION NUMBER: 168,190  
 FILING DATE: 15-MAR-1988  
 APPLICATION NUMBER: 54,369  
 FILING DATE: 26-MAY-1987  
 SEQ ID NO: 1  
 LENGTH: 4383  
 5177307-1



Query Match	27.2%	Score 147.4	DB 6	Length 4383	
Best Local Similarity	67.1%	Pred. No. 3,3e-30			
Matches 261	Conservative 0	Mismatches 111	Indels 17	Gaps 3	

  

QY	82	TCGAAAACAGAAAGCAACACATTTCTTGGAAATATGACTTACCAAAACCAACATGCAGAAAA	141
DB	2942	TGCAAAACACCAAGCCAAACCTTTCCAGAGATTATGTTATATGACTCATCATCTGTAGAAAA	3001
QY	142	GCTTGTAATGAGTGAATTTACTGATGTCATTTGAGCAAAATCTCCAGAGTGCCTA	201
DB	3002	TATTGTATCAAGAAGAAATTTACTGTGTGATGTCATTTGACCAATCTCCAAAGGAAGTGTCTA	3061
QY	202	TGTAATAAGCATGTGTGTGTTGATGAGAAATGACTAAAACAGAGCTGAAATTTTGCT	261
DB	3062	TGCACTAAGCATGTGTATTT---TGACAAAATCTCAAGTGAAGTTAAAGCACTTTGGGT	3118
QY	262	GAGGAAGCAAAAACCTTGGCTGCAGCTTTGCTTGAAGAAGATPATGATTAATTAATTA	321
DB	3119	GAGGAAGCAAAAACCTCTAGAGAAATGTGTGTTGAAGAAGATPATGATGAGATTAATTA	3178
QY	322	GAGATTAGAGAAATTAAGAT-----GCATATACACATATAATTAAGTTCTA	370
DB	3179	TTAAGTAGGTTAATAATAGATTTTGAGTGCAAAAAAACAAATTAATAAGTGTGC	3238
QY	371	CCTTCTTAAGAATGTAGC---TAATGTTGTGTTTAATTTGAGCTTTTGTAGCCTTTTAT	427
DB	3239	CTTTTCTTATAGGTTAGCTGTGATGTGTGTTATGTTATGTTGGCTATATGATGACATTTGA	3288
QY	428	TACACTTAAATTAAGTGTGCACTTCAAT	456
DB	3289	CACATTAATAAGTTGTGACACATCAAT	3327

  

RESULT 10	
US-09-442-631-1	
Sequence 1, Application US/09442631	
Patent No. 6300489	
GENERAL INFORMATION:	
APPLICANT: OH, BOUNG-JUN	
APPLICANT: KO, MOON KYUNG	
APPLICANT: SHIN, BYONGCHUL	
APPLICANT: CHUNG, CHANG HO	
TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND	
TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE	
TITLE OF INVENTION: INCOMPATIBLE INTERACTION	
FILE REFERENCE: 1942/44	
CURRENT APPLICATION NUMBER: US/09/442,631	
CURRENT FILING DATE: 1999-11-18	
NUMBER OF SEQ ID NOS: 4	
SOFTWARE: Patent In Ver. 2.0	
SEQ ID NO 1	
LENGTH: 506	
TYPE: DNA	
ORGANISM: Capsicum annuum	
US-09-442-631-1	

  

Query Match	17.7%	Score 95.6	DB 4	Length 506	
Best Local Similarity	56.9%	Pred. No. 1.2e-16			
Matches 312	Conservative 0	Mismatches 174	Indels 62	Gaps 5	

  

QY	1	ATGGCTCGCTCTGTGCTTCATGGCATTTGCTATCTTGCAAGATGCTTTGTGTGC	60
DB	10	ATGGCTCGTTCCATTTACTTCATGCAATTTCTTGTCTTGCAATGCACTTTTGTGTCT	69
QY	61	TATGAGTGCAGACTAGAGATGCAAAAACAGAAAGCAACATTTCTTGGAATATGATT	120
DB	70	TATGGGTGCAGAGGCAAGAAATTTGCTGTAAAGACTACAAAACCTGTAAATGTCT	129
QY	121	ACCAACCAACATGCAGAAAAGCTGTAT---CAGTGAAGAAATTTACTGATGTCATTTG	177
DB	130	AGTGAACCTCTATGTGCAAAAACCTGTATATGAGAAAGAAATATAGAATGTGTATTGT	189
QY	178	AGCAAAATCTCAGAAAGTCCATATGTACTTAAGCACTGTGTCTTTATATGAAGATGACT	237

```

190 TTCACATCTCTAGCAGATGCTTATGATGAAAGATGTA----- 229
QY 238 AAACAGAGAGCTGAATTTTGGCTGAGAGACAAAACTTTGGCTGAGCTTGTGAA 297
Db 230 -----ATGCTTAAACTCTCGGAAAGAAATTCGTCCT 261
QY 298 GAGAGATTAATGATTAAGTAATTAAGAGATTAGAAATTAAGAGTACAGTATCAACAT 357
262 TGA---AGAAAGAGTTCTTCAAGCAGTAATTAAGTTGATTAATGATTAACTGTCACAA 318
QY 358 AATTAAGT-----TTTACCTTTCTTAAAGTAGTAAAGTAAAGTGTGTTTAAATGGCTTT 413
Db 319 AATTAAATAAAGTGTGCTTCTTCTTAAAGGTAAGTAATTAATGATGTAATCTTGGTGA 378
QY 414 TAGTACCTTTTAAATTAACCTTAAATAAAGTGTGCACTTCAATCTTTGGCAATCTTC 473
379 TAGTACCCATTTGACACATTAATTAATTAAGTTGTGACACATCAATCTTCATGATCTTCT 438
QY 474 ACTAAGTTAATTTGTGTACTTTTATGAATAATGACCTTCCATGCTTGTGGTTAAAAAA 533
Db 439 ATTTAG-----TTGTGTGTTTAAAGAAAGAGATCGTTTACGGTCTTAAAAAAA 494
QY 534 AAAAAAA 541
Db 495 AAAAAAA 502

RESULT 11
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; Prior APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

## IMMEDIATE SOURCE:

CLONE: PT29pt-F18  
US-08-232-463-14

Query Match 8.4%; Score 45.4; DB 1; Length 7218;  
Best Local Similarity 5.5%; Pred. No. 0.0076;  
Matches 19; Conservative 184; Mismatches 140; Indels 0; Gaps 0;

QY 64 GAGGTGCAAGCTAGGATGCAAAACAGAAACACATTTCTCGAATTCGATTACC 123  
DB 1367 RRR 1308  
QY 124 AACACACATGACGAAAAGCTGTATGAGAAATTAAGTATGATGCTTAGAGCAA 183  
DB 1307 RRR 1248  
QY 184 ATCTTCAGAGGTGCTTATGTAAGCAATGTGTTTATGAGACATCACTAAACA 243  
DB 1247 RRR 1188  
QY 244 GGAGCTGAATTTGGCTGAGGAGCAAAACCTTGCTGCTGCTTCTTGAAGAGAG 303  
DB 1187 RRR 1128  
QY 304 ATAATGATTAAGTAAAGTAAAGAAATTAAGATGCAATCAACATTAATAA 363  
DB 1127 RRR 1068  
QY 364 GTTCTACCTTTCTTAAGTATGATGTTTAT 406  
DB 1067 RATTGCAAGCTCCCTGACCTGACGCAAGCTCGGAATTAAT 1025

## RESULT 12

US-09-601-198-63/c  
Sequence 63, Application US/09601198

Patent No. 6531583  
GENERAL INFORMATION:  
APPLICANT: Cassell, Gail H.  
APPLICANT: Chen, Ellison Y.  
APPLICANT: Glaes, Jennifer S.  
APPLICANT: Glaes, John I.  
APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 63  
LENGTH: 1692  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-63

Query Match 7.8%; Score 42; DB 4; Length 1692;  
Best Local Similarity 48.7%; Pred. No. 0.039;  
Matches 114; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 303 GATTAATGATTAATTAAGATTAAGAAATTAAGATGCAATCAACATTAATA 362  
DB 871 GATTAATTAATTAATTAAGTTATTAATTAACAATTTGAATTAATAAAT 812  
QY 363 AGTTTACCTTTCTTAAGAGTACGTAATGTTGTTTAAATGGCTTTAGAGCT 422  
DB 811 CAGCTAATGTTATTAATAATCAATTAATCAATGTTAATGTTCTTAATCTTAA 752  
QY 423 TTATTAAGCTTAATAATAGTGAGCACTTCAATCTTGTGCAATCTTGACTAAGTT 482

DB 751 TTTTGCATCAGTCAAAAAGTGATTTCTTGACCTTTTGTAGCAATTTTATAGCT 692

QY 483 ATTGTGTCCTTTTAAATGAATGACCTTCTATGCTCTTGTAAATAAAAA 536  
DB 691 CTTCAATAATTTGATGTAATTAATAACATCAATATATTTTATATAATGAA 638

## RESULT 13

US-09-801-861-3  
Sequence 3, Application US/09801861

Patent No. 6492154  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
FILE REFERENCE: C1001098  
CURRENT APPLICATION NUMBER: US/09/801,861  
CURRENT FILING DATE: 2001-03-09  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 53332  
TYPE: DNA  
ORGANISM: Human  
US-09-801-861-3

Query Match 7.0%; Score 38; DB 4; Length 53332;  
Best Local Similarity 51.8%; Pred. No. 1.4;  
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 332 GAATTAAGAGTACATCAACATTAATAAGTTCTGCTTCTTAATAAGTAGCTA 391  
DB 31303 GCAATCAGCATATATTTTATATATATTTATATATTTATATATATATATAT 31362  
QY 392 ATGTTGTTTAAATGCTTTAGAGCTTTTATTAACCTTAATTAAGTGGCACT 451  
DB 31363 GTTATATTTTAAATTTTATATATTTATATATTTATATATTTATATATTTATAT 31422  
QY 452 TCAATCCTTTGGCAATCTTGCACTAAGTTTATTTGTGACTTTTA 497  
DB 31423 TTATATATTTTATATATTTATATTTATATTTATATTTTATATATTTTA 31468

## RESULT 14

US-08-472-028A-3  
Sequence 3, Application US/08472028A

Patent No. 5767373  
GENERAL INFORMATION:  
APPLICANT: Ward, Eric R.  
APPLICANT: Volrath, Sandra  
TITLE OF INVENTION: Manipulation of Protoporphyrinogen  
TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: US  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,028A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1748/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1738 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 70..1596 /note= "Arabidopsis protox-2 cDNA"  
OTHER INFORMATION: sequence from pmDC-1"  
US-08-472-028A-3

Query Match 6.9%; Score 37.2; DB 1; Length 1738;  
Best Local Similarity 54.3%; Pred. No. 0.76;  
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 253 ATTTGGCTGAGGAAGCAAAACTTTGGCTGACGCTTTGCTTGAAGAGATATGAT 312  
DB 259 ATTTGGGATGAAGAGCAACACCATGACTGAGGCTGACAGAGTTGGAGTTTACTT 318  
QY 313 AACTAATTAGAGATTAGAAATTAAGATGACATATTAATTAATTTCTACC 372  
DB 319 GATGATCTTGGGCTTCTGAGAACACCAATTTCCATTTCACAGAAAACGGTATATT 378  
QY 373 TTCTTAAAGTGTAGCT 390  
DB 379 GTGGGAATGTGTACTT 396

RESULT 15  
US-08-808-931-3  
Sequence 3, Application US/08808931  
GENERAL INFORMATION:  
APPLICANT: Volzath, Sandra  
APPLICANT: Johnson, Marie  
APPLICANT: Potter, Sharon  
APPLICANT: Ward, Eric  
TITLE OF INVENTION: DNA Molecules Encoding Plant  
TITLE OF INVENTION: Protochlorophyllin Oxidase and Inhibitor-Resistant Mutants  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5939602artis Corporation  
STREET: 520 White Plains Road, P.O. Box 2005  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/808,931  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/012,705  
FILING DATE: 28-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013,612  
FILING DATE: 28-FEB-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,003  
FILING DATE: 21-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Weigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1847  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8587  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1738 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
IMMEDIATE SOURCE:  
CLONE: pmDC-1 (NRRL B-21237)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 70..1596  
OTHER INFORMATION: /product= "Arabidopsis protox-2"  
US-08-808-931-3

Query Match 6.9%; Score 37.2; DB 2; Length 1738;  
Best Local Similarity 54.3%; Pred. No. 0.76;  
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 253 ATTTGGCTGAGGAAGCAAAACTTTGGCTGACGCTTTGCTTGAAGAGATATGAT 312  
DB 259 ATTTGGGATGAAGAGCAACACCATGACTGAGGCTGACAGAGTTGGAGTTTACTT 318  
QY 313 AACTAATTAGAGATTAGAAATTAAGATGACATATTAATTAATTTCTACC 372  
DB 319 GATGATCTTGGGCTTCTGAGAACACCAATTTCCATTTCACAGAAAACGGTATATT 378  
QY 373 TTCTTAAAGTGTAGCT 390  
DB 379 GTGGGAATGTGTACTT 396

Search completed: January 28, 2004, 09:56:13  
Job time: 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 08:44:46 (Search time 1946 Seconds  
(without alignment)) 6756.794 Million cell updates/sec

Title: US-10-072-809b-17  
Perfect score: 541

Sequence: 1 atgcctcgcctcctgtgctt.....tgcgttaaaaaaaaaaaaaa 541

Scoring table: IDENTITY\_NTC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	229.2	42.4	316	13 BU494528 DD-R7-01
2	200.4	37.0	570	12 BU494528 DD-R7-01
3	198.8	36.7	570	12 BU494528 DD-R7-01
4	198.8	36.7	570	12 BU494528 DD-R7-01

5	196.4	36.3	533	12 BU494528 DD-R7-01
6	194	35.9	541	12 BU494528 DD-R7-01
7	194	35.9	541	12 BU494528 DD-R7-01
8	189.8	35.1	522	9 A1485044 EST243324
9	189.8	35.1	522	9 A1485044 EST243324
10	189.8	35.1	522	9 A1485044 EST243324
11	189.8	35.1	522	9 A1485044 EST243324
12	189.8	35.1	522	9 A1485044 EST243324
13	189.8	35.1	522	9 A1485044 EST243324
14	189.8	35.1	522	9 A1485044 EST243324
15	188.2	34.8	500	9 A1485044 EST243324
16	188.2	34.8	500	9 A1485044 EST243324
17	186.6	34.5	562	9 A1485044 EST243324
18	186.6	34.5	562	9 A1485044 EST243324
19	183.8	34.0	567	9 A1485044 EST243324
20	183.8	34.0	567	9 A1485044 EST243324
21	182.8	33.8	561	12 BU494528 DD-R7-01
22	182.8	33.8	561	12 BU494528 DD-R7-01
23	181.8	33.6	419	9 A1485044 EST243324
24	179.4	33.2	563	9 A1485044 EST243324
25	176.8	32.7	574	12 BU494528 DD-R7-01
26	170	31.4	378	9 A1485044 EST243324
27	170	31.4	378	9 A1485044 EST243324
28	169.2	31.3	435	10 A1485044 EST243324
29	167	30.9	481	10 A1485044 EST243324
30	167	30.9	481	10 A1485044 EST243324
31	166.4	30.8	513	12 BU494528 DD-R7-01
32	163.6	30.2	491	12 BU494528 DD-R7-01
33	162.6	30.1	573	14 A1485044 EST243324
34	160.8	29.7	403	12 BU494528 DD-R7-01
35	160	29.6	500	12 BU494528 DD-R7-01
36	158.2	29.2	505	9 A1485044 EST243324
37	156.4	28.9	421	9 A1485044 EST243324
38	154.8	28.6	536	12 BU494528 DD-R7-01
39	152.2	28.1	589	14 A1485044 EST243324
40	148.6	27.5	333	9 A1485044 EST243324
41	145.8	27.0	476	10 A1485044 EST243324
42	137.8	25.5	537	12 BU494528 DD-R7-01
43	134.6	24.9	532	12 BU494528 DD-R7-01
44	133	24.6	571	14 A1485044 EST243324
45	132.2	24.4	603	9 A1485044 EST243324

## ALIGNMENTS

RESULT 1  
BU494528 316 bp mRNA linear EST 09-JAN-2003  
LOCUS DD-R7-01 DD-R7-derived M. sexta/T. notatus-responsive library  
DEFINITION Nicotiana attenuata cDNA clone cys8.4.3, similar to Solanaceae mRNA  
for flower-specific chitinase (Z11748.1), mRNA sequence.

ACCESSION BU494528  
VERSION BU494528.1 GI:27552699  
KEYWORDS EST.

SOURCE Nicotiana attenuata  
ORGANISM Nicotiana attenuata

REFERENCE  
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; Lamiales; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 316)  
TITLE Voelckel, C. and Baldwin, I.T.  
Detecting herbivore-specific transcriptional responses in plants  
with multiple DDRT-PCR and subtractive library procedures  
Unpublished

JOURNAL COMMENT  
Contact: Gase X  
Department of Molecular Ecology  
Max Planck Institute for Chemical Ecology  
Winzlerstr. 10, Beutenberg Campus, 07745 Jena, Germany  
Tel: 49/3641/571121  
Fax: 49/3641/571102  
Email: gase@ice.mpg.de  
Seq primer: M13 Forward (-20)

POLYA=Yes.

FEATURES  
source

Location/Qualifiers  
1. 316  
/organism="Nicotiana attenuata"  
/mol\_type="mRNA"  
/db\_xref="taxon:49451"  
/clone="cv84.4"  
/tissue\_type="shoot"  
/dev\_stage="rosette"  
/lab\_host="E. coli"  
/clone\_lib="Dd-R7-derived M. sexta/T. notatus-responsive library"  
/note="Vector: PCR 2.1-TOPO; Site 1: EcoR I; Site 2: EcoR I; DdR7-PCR with arbitrary primer R7 (Nucleic Acid Research 21, 4272-4280, 1993)"

BASE COUNT  
ORIGIN  
97 a 44 c 66 g 109 t

Query Match 42.4%; Score 229.2; DB 13; Length 316;  
Best Local Similarity 91.4%; Fred. No. 2.4e-31;  
Matches 265; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

229 AACAGGAGCTGAAATTTGGCTGAGAGCAAAACCTTGGCTGACGCTTGGCTTGAAG 298  
4 ATACGAGAGCTGAACTTTAGCCGAGAGGCAACACCTTGGCTGAAAGCTTGGCTTGAAG 63  
299 AAGAGATATGATTAATTAATTAAGATTAGAGAAATTAAGATGCAATACACATA 358  
64 AAGAGATATGATTAATTAATTAAGATTAGAGAAATTAAGATGCAATACACATA 123  
359 ATAAAGTTTCACTTCTTAATTAAGTTAGCTTAATGTTTGAATTTGCTTTAGTA 418  
124 ATAAAG-TGCTGCTTTCTTAATTAAGTTAGCTTAATGTTTGAATTTGCTTTAGTA 182  
419 GCTTTTATTAACCTTTAAATTAAGTTGAGCACTTCAATCTTTGTCATCTTGAATTA 478  
133 GCCGTTTGTACCTTTAAATTAAGTTGAGCACTTCAATCTTTGTCATCTTGAATTA 242  
479 GTTATTTTGTACTTTTAATTAAGTTGAGCACTTCAATCTTTGTCATCTTGAATTA 528  
243 GTTATTTTGTACTTTTAATTAAGTTGAGCACTTCAATCTTTGTCATCTTGAATTA 291

RESULT 2  
BI929149 570 bp mRNA linear EST 18-OCT-2001  
LOCUS EST149038 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA

DEFINITION  
EST149038 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA  
clone CTOB27L16 5' end, mRNA sequence.

ACCESSION  
BI929149  
VERSION  
BI929149.1 GI:16242157  
KEYWORDS  
EST.

SOURCE  
ORGANISM  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
AUTHORS  
van der Hoeven, R.S., Bezzerides, J.L., Karanycheva, S.A., Tsai, J.,  
Uteirback, T., Van Aken, S., Roming, C.M., Niernman, W., Fraser, C.M.,  
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)

TITLE  
JOURNAL  
COMMENT  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3

FEATURES  
source  
1. 570  
/organism="Lycopersicon esculentum"

/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CTOB27L16"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/note="Vector: pBluescript SK(-); Site 1: EcoR; Site 2: XhoI; supplier: Cornell University; sequencing: the institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT  
ORIGIN  
179 a 81 c 119 g 191 t

Query Match 37.0%; Score 200.4; DB 12; Length 570;  
Best Local Similarity 68.5%; Fred. No. 3e-26;  
Matches 370; Conservative 0; Mismatches 151; Indels 19; Gaps 6;

1 ATGCTTCGCTCTTGTGCTTCAATGAGCATTTGCTTGAAGAGATGCTCTTGTGCTC 60  
34 ATGGCTCGTTCATTTGCTTCAATGAGCATTTGCTTGAAGAGATGCTCTTGTGCTC 93  
61 TATGAGGTGCAATGAG 117  
94 TCTGAGGTGCAATGAG 153  
118 ATTACCAACCAATGAG 177  
154 TTACCGATTCATCTGAG 213  
178 AGCAAAATCTGACT 237  
214 AGCAAAATCTGACT 273  
238 AAAACAGAGCTGAAATTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297  
274 GAAGTTAAAC--ACCTTGGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330  
298 GAAGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 351  
331 GAAGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 390  
352 ACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 411  
391 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 450  
412 TTATGAGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 468  
451 TTATGAGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 508  
469 CTTCAGTAAGTTTATTTGTGA--CTTTTAAATGAATGACCTTCTATGCTTTGAT 526  
509 CTTCAGTAAGTTTATTTGTGA--CTTTTAAATGAATGACCTTCTATGCTTTGAT 568

RESULT 3  
BI927088 570 bp mRNA linear EST 18-OCT-2001  
LOCUS EST146977 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA

DEFINITION  
EST146977 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA  
clone CTOB2101 5' end, mRNA sequence.

ACCESSION  
BI927088  
VERSION  
BI927088.1 GI:16236262  
KEYWORDS  
EST.

SOURCE  
ORGANISM  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE  
AUTHORS  
van der Hoeven, R.S., Bezzerides, J.L., Karanycheva, S.A., Tsai, J.,  
Uteirback, T., Van Aken, S., Roming, C.M., Niernman, W., Fraser, C.M.,  
1 (bases 1 to 570)

Thu Jan 29 11:57:50 2004

us-10-072-809b-17.rst

Page 3

TITLE  
Journal  
COMMENT  
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute  
Seq primer: T3.

FEATURES  
SOURCE

1..570  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLOB2101"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT  
180 a 81 c 119 g 190 t

ORIGIN

Query Match  
Best Local Similarity 68.3%; Score 198.8; DB 12; Length 570;  
Matches 369; Conservative 0; Mismatches 152; Indels 19; Gaps 6;

1 ATGCTCGCTCTCTGCTTCATGGCATTTGCTATCTTGGCAGAGATCTCTTGTGCC 60  
34 ATGGCTCGCTCTCTGCTTCATGGCATTTGCTATCTTGGCAGAGATCTCTTGTGCC 93  
61 TATGAGTGCAGCTAG 117  
94 TCTGAGTGCAGCTAG 153  
118 ATTACCAACACCAATGCAAGAAAGCTGTATCAAGTGAATTTACTGATGCTATTGT 177  
154 TTACCGATTCATCGTGAAGAGCTGTGTCAAGAGAGAGAGAGAGAGAGAGAGAG 213  
178 AGCAAAATCTCAG 237  
214 AGCAAAATCTCAG 273  
238 AAAACAG 297  
274 GAAGTTAAAC--AACTTGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330  
298 GAAGAGTAATGATTAATTA-----GAGATTGAAGAAATTAAGATGAGATGCT 351  
331 GAAGAGATGATTAATTAATTAAGTGAATTAAGATGATTAAGATGATTAAGATG 390  
352 ACACATTAATTAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 411  
391 AAAATTAATTAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 450  
412 TTTAGAGAGCTTTTATTAATTAATTAAGTGCAGCTTAAT--CTTTGTCAGAT 468  
451 TTTAGAGAGCTTTTATTAATTAATTAAGTGCAGCTTAAT--CTTTGTCAGAT 508  
469 CTGCACTAAGTTTATTTGTGTA--CTTTTATGAAGAAATGACCTTCTTAAGTCTTGT 526  
509 CTGCACTAAGTTTATTTGTGTA--CTTTTATGAAGAAATGACCTTCTTAAGTCTTGT 568

RESULT 4  
B1928623

LOCUS  
EST548512 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA  
clone CLOB26K7 5' end, mRNA sequence.  
B1928623  
ACCESSION  
B1928623.1 GI:16240393  
VERSION  
B1928623.1  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Lycopersicon esculentum (tomato)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 570)  
van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,  
Utterback, T., Van Aken, S., Ronning, C.M., Niernman, W., Fraser, C.M.,  
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)  
Unpublished  
Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute  
Seq primer: T3.

REFERENCE  
AUTHORS  
TITLE  
Journal  
COMMENT  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute  
Seq primer: T3.

FEATURES  
SOURCE

1..570  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLOB26K7"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/clone\_lib="tomato flower, 3 - 8 mm buds"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

BASE COUNT

180 a 81 c 119 g 190 t  
Query Match  
Best Local Similarity 68.3%; Score 198.8; DB 12; Length 570;  
Matches 369; Conservative 0; Mismatches 152; Indels 19; Gaps 6;

1 ATGCTCGCTCTCTGCTTCATGGCATTTGCTATCTTGGCAGAGATCTCTTGTGCC 60  
34 ATGGCTCGCTCTCTGCTTCATGGCATTTGCTATCTTGGCAGAGATCTCTTGTGCC 93  
61 TATGAGTGCAGCTAG 117  
94 TCTGAGTGCAGCTAG 153  
118 ATTACCAACACCAATGCAAGAAAGCTGTATCAAGTGAATTTACTGATGCTATTGT 177  
154 TTACCGATTCATCGTGAAGAGCTGTGTCAAGAGAGAGAGAGAGAGAGAGAGAG 213  
178 AGCAAAATCTCAG 237  
214 AGCAAAATCTCAG 273  
238 AAAACAG 297  
274 GAAGTTAAAC--AACTTGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330  
298 GAAGAGTAATGATTAATTA-----GAGATTGAAGAAATTAAGATGAGATGCT 351  
331 GAAGAGATGATTAATTAATTAAGTGAATTAAGATGATTAAGATGATTAAGATG 390  
352 ACACATTAATTAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 411

Db	391	AAATTTAAAGTATTGCTCTTTAGTAAAGGTACTCTTATATGCTGTTATTG	969C	450
QY	412	TTTAGTAGCCTTTTATTAACACTTAAATAGTGTGCACTTCAAT--CCTTTGGCAAT		468
Db	451	TTTAGTAGCATTATTAACA--TTAAATTAATTTTGACAGTCATTATCTTTGGATT		508
QY	469	CTTGACCTAGTTTATTGTGTA--CTTTATGAAATACCTTTATGTCCTTTGGTT		526
Db	509	CTTGATCAACTTAATCACTGTAACTTTAATGAAAAATATATGACATAGTCTTTAATT		568

	EST	Accession
RESULT 5	B1927975	
LOCUS	B1927975	
DEFINITION	B1927975 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA	
ACCESSION	B1927975	
VERSION	B1927975.1	
KEYWORDS	EST.	
SOURCE	Lycopersicon esculentum (tomato)	
ORGANISM	Lycopersicon esculentum	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (Pages 1 to 533)	Bezzerides, J.L., Karayancheva, S.A., Tsai, J., van der Hoeven, R.S., Romling, C.M., Niemann, W., Fraser, C.M., Utecht-Back, T., Gan Akeni, J., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.	Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)	Unpublished	Contact: CUGI

**FEATURES**  
**Source**      **Location/Qualifiers**  
 1..553      **1..553**

```

BASE COUNT      168 a      72 c      113 g      180 t
ORIGIN
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/c1one="cTOB24E21"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/c1one_1db="tomato flower, 3 - 8 mm buds"
/note="vector: plnuescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research. Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

	Query Match	Match Similarity	Best Local	Matches	Conservative	Score	Pred. No.	DB	Length	Indels	Gaps
Qy	5	CTCGCTCTTGTCCTTCATGCAATTCGATCTTGGCAAGATGCTCTTTGTGCTCATG	64			196.41	12	533			
Db	1	CTCGTTCATTGTCCTTCATGCACTATAGTCTTGGCAATGCTCTTTTTCCTCG	60			68.33	1,6e-25				
Qy	65	AGGTGCAAGCTAGAGA---ATGCAAAACAGAAAGCAACATTCCTCGAATATGCAATTA	121								
Db	61	AGGTGCAAGCTCAGACAGATGCAAAATCAACAGCAAACTTCAAGGATTTAGCTTTA	120								
Qy	122	CCAAACCAACATGCAAGAAAGCTTGTATCAGTAGAATTTACTGATGATCTATTTGACA	181								
Db	121	CGATATGATCGTAGAAGAGCTGTGTCTCACAAGAGATTTACAGTGGACATTTGACA	180								

Oy	182	AAATCCCAAGAGTGCCTTAATGACTAAGCCATGTCGTTGATGGAAGAATGACTAAA	244
Oy	181	AACTCAAAAGAAAGTGCCTTAATGACTAAGTGTGTGATTTGAAAAAGATTCAAAATGAAG	240
Oy	242	CAGAGCTGAATTTTGCTGAGAGACAAAACCTTGGCTGCAGCTTGTGTTGAAGAAG	301
Db	241	TTAAATAC---AACCTTGTTGGGAGACAAAACCTTAAGTGAACGTGTGTGAAGAAG	297
Oy	302	AGATTAATGAGTAATCTAATTA-----GAGATTGAGAAATTAAGAATGACATGACAC	355
Db	298	AGATTGTGATGGAGTAAATTAATAGTGAATTAATTAAGAATTGAGTGTCAAAATCAAAA	357
Oy	356	ATAATTAAGTTTCTAACCCTTTAAAGTGAAGTAAATGTTGCTTTTAATGTGCTTTTA	415
Db	358	TTAATTAAGTATGTCCTTTTGTGTAAGAGGATGCTTTAATGTGTGTATATGGGCTTTA	417
Oy	416	GTACCCCTTTTATTAACACTTAAATTAAGTGTGCACTTCAAT---CCTTGTGCATCTTG	472
Db	418	GTACCCACTTTGATACAA---TTAAATAAGTTTGAATGATCAATTAATCCTTTTGGATTCTTG	475
Oy	473	CACTAAGTTAATTTGTGTA---CCTTTATATGAATAAGACCTCTAATGCTTGTGGTT	526
Db	476	TATCAACTTAATTTAGTGTAACTTTTAATGAATAATGATCGACATATGGCTTTAATT	531

RESULT 6  
BI929019  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BI929019  
EST548908 tomato flower, 3-8 mm buds  
clone cT0827M14.5' end, mRNA sequence.  
BI929019  
BI929019.1 GI:16241728  
EST.  
Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (baes 1 to 541)  
van der Hoeven, R.S., Bezzerides, J.L., Karamcheva, S.A., Tsai, J.,  
Utterback, T., Van Aken, S., Rouning, C.M., Nierman, W., Fraser, C.M.,  
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)  
Unpublished  
Contact: CUGI

	106	October 1997	Genome, Email: <a href="http://www.genome.clemson.edu/orders/index.html">http://www.genome.clemson.edu/orders/index.html</a> This clone is available through the Clemson University Genomics Institute
FEATURES			
source			
	Seq primer: T3.		
	Location/Qualifiers		
	1..541		
	/organism="Lycopersicon esculentum"		
	/mol_type="cDNA"		
	/cultivar="TA496"		
	/db_xref="taxon:4081"		
	/clone="CLOB27M14"		
	/tissue_type="flower"		
	/dev_stage="3-8mm buds"		
	/clone_1lb="10cmato flower, 3 - 8 mm buds"		
	/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University sequencing. The institute for Genomic Research. Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."		
BASE COUNT	170 a	78 c	115 g 178 t
ORIGIN			
Query Match	35.9%	Score 194;	DB 12; Length 541;

Best Local Similarity 68.2%; Pred. No. 4.3e-25;  
Matches 348; Conservative 0; Mismatches 145; Indels 17; Gaps 5;

QY 1 ATGGCTGCTCTCTGCTTCTCATGAGATTTGCTATCTTGGCAAGATGCTCTTTGTGCC 60  
DB 34 ATGGCTGCTCTCTGCTTCTCATGAGATTTGCTATCTTGGCAAGATGCTCTTTGTGCC 93  
QY 61 TATGAGTGCAGCTGAGAG--ATGCAAAACAGAAAGCAACATTTCTCGAATATGC 117  
DB 94 TGTGAGGTGCAGCTGAGAGATGAGCAATCAACAGCAACCTTCAAGGATTTATGC 153  
QY 118 ATTACCAACCAACCATGACAGAAAGCTTATCACTGAGAAATTTACTGATGCTATGT 177  
DB 154 TTACCGATTCATCGTGTAGAAAGCTTGTCTCAGAGAGATTTACAGGTGATGCT 213  
QY 178 AGCAAAATCTCAGAAAGTGTCTTACTTACGATGTGTGTGTGTGTGTGTGTGTGT 237  
DB 214 AGCAAACTCCAAAGAAAGTGTCTTACTTACGATGTGTGTGTGTGTGTGTGTGTGT 273  
QY 238 AAAACAGAGCTGAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 297  
DB 274 GAAATTTAAAC--AACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 330  
QY 298 GAAAGATTAATGATTAATTA--GAGATTAGAGAAATTAAGATGACATATC 351  
DB 331 GAAAGATTTGTATGATTAATTAATTAAGATTAATTAAGATTTGAGTGTCAATTC 390  
QY 352 AACATTAATTAATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 411  
DB 391 AAAATTAATTAATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 450  
QY 412 TTATAGACCTTTTATTAACCTTTAAATAGTGTGCACTTCAAT--CCTTTGTCAAT 468  
DB 451 TTATAGACCTTTTATTAAC--TTAAATAGTGTGCACTTCAATTTCTTTGTGAT 508  
QY 469 CTTCGACTAGTTTATTTGTGTACTTTAA 498  
DB 509 CTGTATCACTTAATCACTGATTAAGTTTAA 538

RESULT 7  
BI929156 541 bp mRNA linear EST 18-OCT-2001  
LOCUS EST249045 tomato flower, 3-8 mm buds Lycopersicon esculentum cDNA  
DEFINITION clone cTOB27N10 5' end, mRNA sequence.  
ACCESSION BI929156  
VERSION BI929156.1 GI:16242173  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 541)  
van der Hoeven, R.S., Bezzerides, J.L., Karanicheva, S.A., Tsai, J.,  
Ullrich, T., Van Aken, S., Roming, C.M., Nierman, W., Fraser, C.M.,  
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)  
Unpublished  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3

FEATURES  
source  
Location/Qualifiers  
1..541  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/culti\_var="TA936"  
/db\_xref="taxon:4081"

/clone="cTOB27N10"  
/tissue\_type="flower"  
/dev\_stage="3-8mm buds"  
/clone\_1ib="tomato flower, 3-8 mm buds"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; supplier: PBI Research, Inc.; sequencing: The  
Institute for Genomic Research, Flower buds and flowers  
were taken from greenhouse plants (4-8 wks old, TA936).  
They were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."  
BASE COUNT 170 a 78 c 115 g 178 t  
ORIGIN

Query Match 35.9%; Score 194; DB 12; Length 541;  
Best Local Similarity 68.2%; Pred. No. 4.3e-25;  
Matches 348; Conservative 0; Mismatches 145; Indels 17; Gaps 5;

QY 1 ATGGCTGCTCTCTGCTTCTCATGAGATTTGCTATCTTGGCAAGATGCTCTTTGTGCC 60  
DB 34 ATGGCTGCTCTCTGCTTCTCATGAGATTTGCTATCTTGGCAAGATGCTCTTTGTGCC 93  
QY 61 TATGAGTGCAGCTGAGAG--ATGCAAAACAGAAAGCAACATTTCTCGAATATGC 117  
DB 94 TGTGAGGTGCAGCTGAGAGATGAGCAATCAACAGCAACCTTCAAGGATTTATGC 153  
QY 118 ATTACCAACCAACCATGACAGAAAGCTTATCACTGAGAAATTTACTGATGCTATGT 177  
DB 154 TTACCGATTCATCGTGTAGAAAGCTTGTGTCTCAGAGAGATTTACAGGTGATGCT 213  
QY 178 AGCAAAATCTCAGAAAGTGTCTTACTTACGATGTGTGTGTGTGTGTGTGTGTGT 237  
DB 214 AGCAAACTCCAAAGAAAGTGTCTTACTTACGATGTGTGTGTGTGTGTGTGTGTGT 273  
QY 238 AAAACAGAGCTGAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 297  
DB 274 GAAATTTAAAC--AACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 330  
QY 298 GAAAGATTAATGATTAATTA--GAGATTAGAGAAATTAAGATGACATATC 351  
DB 331 GAAAGATTTGTATGATTAATTAATTAAGATTAATTAAGATTTGAGTGTCAATTC 390  
QY 352 AACATTAATTAATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 411  
DB 391 AAAATTAATTAATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 450  
QY 412 TTATAGACCTTTTATTAACCTTTAAATAGTGTGCACTTCAAT--CCTTTGTCAAT 468  
DB 451 TTATAGACCTTTTATTAAC--TTAAATAGTGTGCACTTCAATTTCTTTGTGAT 508  
QY 469 CTTCGACTAGTTTATTTGTGTACTTTAA 498  
DB 509 CTGTATCACTTAATCACTGATTAAGTTTAA 538

RESULT 8  
AI485044 522 bp mRNA linear EST 18-MAY-2001  
LOCUS EST243324 tomato ovary, TAMO Lycopersicon esculentum cDNA clone  
DEFINITION clone D2B28, mRNA sequence.  
ACCESSION AI485044  
VERSION AI485044.1 GI:4380415  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 522)  
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,  
Li, F., Upton, J., Roming, C.M., Craven, M.B., Fujii, C.Y., Bowman,  
C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley,  
S.D. and Giovannoni, J.J.  
Generation of ESTs from tomato carpel tissue

FEATURES  
AUTHORS  
TITLE  
Location/Qualifiers  
1..522  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/culti\_var="TA936"  
/db\_xref="taxon:4081"



JOURNAL Unpublished  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES Location/Qualifiers

Source

```
1..522
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLED218"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MRF"
/clone_id="tomato ovary, TAMC"
/note="Vector: Bluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST Library. Oligodt-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
```

BASE COUNT 166 a 77 c 102 g 177 t

ORIGIN

Query Match 35.1%; Score 189.8; DB 9; Length 522;  
Best Local Similarity 68.9%; Pred. No. 2.5e-24;  
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

```
QY 1 ATGGCTGCTCTTGTGCTTCATGCGCATTTGCTATCTTGGCAGAGATGCTTTGTTGCC 60
DB 27 ATGGCTGCTCTTGTGCTTCATGCGCATTTGCTATCTTGGCAGAGATGCTTTGTTACC 86
QY 61 TATGAGTGCAGAGCTAGAGAA---TGCAAAACAGAAACACATTTCTCGAATATGC 117
DB 67 TATGAGTGCAGAGCTAGAGAA---TGCAAAACAGAAACACATTTCTCGAATATGC 146
QY 118 ATTACCAACACACATGAGAAAGCTTGTATCGATGAGAAATTTACTGATGTCATTGT 177
DB 147 TTTATGACTCATCATGAGAAATTTGTATCGAAGAAATTTACTGATGTCATTGT 206
QY 178 AGCAAAATCTCAGAAAGTGCCTATGTAAGCATGATGTTGATGAGAGATGACT 237
DB 207 AGCAAAATCTCAGAAAGTGCCTATGTAAGCATGATGTTGATGAGAGATGACT 263
QY 238 AAAACAGAGCTGAATTTTGGCTGAGAGACAAAACCTTTGGCTGAGCTTTGCTTGA 297
DB 264 AGTGAAGTTAAAGCACTTTGGCTGAGAGACAAAACCTTTGAGTGAAGTTGCTTGA 323
QY 298 GAAAGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 346
DB 324 GAAAGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 383
QY 347 GTATCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403
DB 384 AAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 443
QY 404 AATTGGCTTTAGTACCTTTTATTAACCTTTAATTAATTAATTAATTAATTAATTAAT 456
DB 444 TATTGGCTTTAGTACCTTTTATTAACCTTTAATTAATTAATTAATTAATTAATTAAT 496
```

RESULT 9  
AI487398 522 bp mRNA linear EST 18-MAY-2001  
LOCUS EST143720 tomato ovary, TAMC Lycopersicon esculentum cDNA clone  
DEFINITION cLED13821, mRNA sequence.  
ACCESSION AI487398  
VERSION AI487398.1 GI:4382769  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES  
Source

```
1..522
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLED13821"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MRF"
/clone_id="tomato ovary, TAMC"
/note="Vector: Bluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST Library. Oligodt-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
```

BASE COUNT 165 a 75 c 104 g 178 t

ORIGIN

Query Match 35.1%; Score 189.8; DB 9; Length 522;  
Best Local Similarity 68.9%; Pred. No. 2.5e-24;  
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

```
QY 1 ATGGCTGCTCTTGTGCTTCATGCGCATTTGCTATCTTGGCAGAGATGCTTTGTTGCC 60
DB 27 ATGGCTGCTCTTGTGCTTCATGCGCATTTGCTATCTTGGCAGAGATGCTTTGTTACC 86
QY 61 TATGAGTGCAGAGCTAGAGAA---TGCAAAACAGAAACACATTTCTCGAATATGC 117
DB 67 TATGAGTGCAGAGCTAGAGAA---TGCAAAACAGAAACACATTTCTCGAATATGC 146
QY 118 ATTACCAACACACATGAGAAAGCTTGTATCGATGAGAAATTTACTGATGTCATTGT 177
DB 147 TTTATGACTCATCATGAGAAATTTGTATCGAAGAAATTTACTGATGTCATTGT 206
QY 178 AGCAAAATCTCAGAAAGTGCCTATGTAAGCATGATGTTGATGAGAGATGACT 237
DB 207 AGCAAAATCTCAGAAAGTGCCTATGTAAGCATGATGTTGATGAGAGATGACT 263
QY 238 AAAACAGAGCTGAATTTTGGCTGAGAGACAAAACCTTTGGCTGAGCTTTGCTTGA 297
DB 264 AGTGAAGTTAAAGCACTTTGGCTGAGAGACAAAACCTTTGAGTGAAGTTGCTTGA 323
QY 298 GAAAGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 346
DB 324 GAAAGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 383
QY 347 GTATCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403
DB 384 AAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 443
QY 404 AATTGGCTTTAGTACCTTTTATTAACCTTTAATTAATTAATTAATTAATTAATTAAT 456
DB 444 TATTGGCTTTAGTACCTTTTATTAACCTTTAATTAATTAATTAATTAATTAATTAAT 496
```

RESULT 10  
AI489818 522 bp mRNA linear EST 18-MAY-2001  
LOCUS AI489818

DEFINITION EST248157 tomato ovary, TAMU Lycopersicon esculentum cDNA clone  
 CLD15G20, mRNA sequence.  
 ACCESSION A1489818  
 VERSION A1489818.1 GI:4385189  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 522)  
 Alcala, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E.,  
 Liang, F., Upton, J., Romling, C. M., Craven, M. B., Fujii, C. Y., Bowman,  
 C. L., Niernman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley,  
 S. D. and Giovannoni, J.  
 Generation of ESTs from tomato carpel tissue  
 Unpublished  
 CONTACT: CUGI  
 JOURNAL  
 COMMENT  
 FEATURES  
 source  
 1..522  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLD15G20"  
 /tissue\_type="carpel"  
 /dev\_stage="5 days pre-anthesis to 5 days post-anthesis"  
 /lab\_host="XLI-Blue MRF"  
 /clone\_1lb="tomato ovary, TAMU"  
 /note="Vector: Bluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLED - Tomato Carpel EST Library. Oligodt-primed and  
 directionally cloned cDNA in vector lambda ZAP II with 5'  
 and 3' ends located at the EcoRI and XhoI sites,  
 respectively."  
 BASE COUNT 165 a 75 c 104 g 178 t  
 ORIGIN  
 Query Match 35.1%; Score 189.8; DB 9; Length 522;  
 Best Local Similarity 68.9%; Pred. No. 2.5e-24;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;  
 1 ATGGCTGCTCTTGTGCTTATGAGCAATTTGCTATCTTGGCAGAGAGCTCTTTGTTGCC 60  
 27 ATGGCTGCTCTTGTGCTTATGAGCAATTTGCTATCTTGGCAGAGAGCTCTTTGTTACC 86  
 61 TATGAGGTGCAAGCTAGAGAA--TGCAAAACAGAAAGCAACATTTCCGGAATATGC 117  
 87 TATGAGGTGCAAGCTAGAGAAATTTGCAAGCAACCAAACTTTCCGAGATTATGT 146  
 118 ATTACCAACCAACCATGAGAAAGCTTGTATCAGTGAATTTACTGATGTCATTGT 177  
 147 TTTATGACTCATCTATGTAGAAATTTGTATCAAGAAATTTACTGAGTGACATTGT 206  
 178 AGCAAAATCTCAGAGAGGTGCTATGTATCAAGCAATGCTGTTGATGAGAAATGACT 237  
 207 AGCAAACTCCAAAGAGAGTGTATGACCTAAGCCATGTGATTT--TGACAAATCTCA 263  
 228 AAAACAGAGCTGAATTTTGGCTGAGAGCAAAATTTTGGCTGAGCTTTGCTTGA 297  
 264 AGTGAAGTTAAAGCACTTTGGGTGAGAGCAAAATCTTAAGTGAAGTTGCTTGA 323  
 298 GAAGGATTAATGATGAGATTAATTAAGTGAAGTTTGAAGTGAAGTGAAGTGA 346  
 324 GAAGGATTAATGATGAGATTAATTAAGTGAAGTTTGAAGTGAAGTGAAGTGA 383  
 347 GTATCACAATTAATTAAGTTTCTAAGCTTTCTTAAAGTGAAGTGAAGTGAAGTGA 403  
 384 AAAACAAATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 443

RESULT 11  
 AM217379  
 LOCUS  
 DEFINITION EST296063 tomato flower buds 0-3 mm, Cornell University  
 Lycopersicon esculentum cDNA clone cTOA1K13 similar to  
 gamma-tubulin-like protein/acidic protein precursor, mRNA sequence.  
 1 (bases 1 to 551)  
 van der Hoeven, R. S., Bezzerides, J. L., Matern, A. L., Holt, I. E., Liang,  
 F., Upton, J., Hansen, T., Craven, M. B., Bowman, C. L., Ahn, S., Romling,  
 C. M., Fraser, C. M., Martin, G. B., Giovannoni, J. J. and Tanksley, S. D.  
 Generation of ESTs from tomato flower tissue  
 Unpublished  
 CONTACT: CUGI  
 JOURNAL  
 COMMENT  
 FEATURES  
 source  
 1..551  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cTOA1K13"  
 /tissue\_type="flower"  
 /dev\_stage="0-3mm buds"  
 /clone\_1lb="tomato flower buds 0-3 mm, Cornell University"  
 /note="Vector: Bluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; supplier: Tanksley; Flower buds and flowers were  
 taken from greenhouse plants (4-8 wks old, TA496). They  
 were immediately frozen in liquid nitrogen and then  
 size-separated while remaining frozen."  
 BASE COUNT 176 a 79 c 107 g 189 t  
 ORIGIN  
 Query Match 35.1%; Score 189.8; DB 9; Length 551;  
 Best Local Similarity 68.9%; Pred. No. 2.4e-24;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;  
 1 ATGGCTGCTCTTGTGCTTATGAGCAATTTGCTATCTTGGCAGAGAGCTCTTTGTTGCC 60  
 27 ATGGCTGCTCTTGTGCTTATGAGCAATTTGCTATCTTGGCAGAGAGCTCTTTGTTACC 86  
 61 TATGAGGTGCAAGCTAGAGAA--TGCAAAACAGAAAGCAACATTTCCGGAATATGC 117  
 87 TATGAGGTGCAAGCTAGAGAAATTTGCAAGCAACCAAACTTTCCGAGATTATGT 146  
 118 ATTACCAACCAACCATGAGAAAGCTTGTATCAGTGAATTTACTGATGTCATTGT 177  
 147 TTTATGACTCATCTATGTAGAAATTTGTATCAAGAAATTTACTGAGTGACATTGT 206  
 178 AGCAAAATCTCAGAGAGGTGCTATGTATCAAGCAATGCTGTTGATGAGAAATGACT 237  
 207 AGCAAACTCCAAAGAGAGTGTATGACCTAAGCCATGTGATTT--TGACAAATCTCA 263  
 228 AAAACAGAGCTGAATTTTGGCTGAGAGCAAAATTTTGGCTGAGCTTTGCTTGA 297  
 264 AGTGAAGTTAAAGCACTTTGGGTGAGAGCAAAATCTTAAGTGAAGTTGCTTGA 323

QY 298 GAAGAGATATGATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346  
 DB 324 GAAGAGATTAATGATGATTAATTAATTAAGATTAGAGAAATTAAGATTTTGAAGTCAAAA 383  
 QY 347 GATACACATTAATTAAGTTTCTTACCTTTCTTAAAGTAC---TAATGTGTGTTT 403  
 DB 384 AAACAAATTAATTAAGTTTCTTACCTTTCTTAAAGTACCTTGTGTGTGTAG 443  
 QY 404 AATTGCTTTAGTACCTTTTATTAACCTTTAATTAAGTGGCACTTCAT 456  
 DB 444 TATTGGCTTATGATGACCATTTGACACATTAATTAAGTTGTGACACATCAT 496

RESULT 12  
 A1897456 562 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST266899 tomato ovary, TAMU Lycopersicon esculentum cDNA clone  
 DEFINITION CLEDD27F13, mRNA sequence.  
 ACCESSION A1897456  
 VERSION A1897456.1 GI:5603358  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanales; Solanum; Lycopersicon.  
 1 (bases 1 to 562)  
 Alcala,J., Vredalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,  
 Liang,F., Upton,J., Roming,C.M., Craven,M.B., Fujii,C.Y., Bowman,  
 C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,  
 S.D. and Giovannoni,J.  
 Generation of ESTs from tomato carpel tissue  
 Unpublished  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY Genomics Institute  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1..562  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEDD27F13"  
 /tissue\_type="carpel"  
 /dev\_stage="5 days pre-anthesis to 5 days post-anthesis"  
 /lab\_host="X11-Blue MRF"  
 /clone\_1lb="tomato ovary, TAMU"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; CLEDD - Tomato Carpel EST Library. Oligodt-primed and  
 directionally cloned cDNA in vector lambda ZAP II with 5'  
 and 3' ends located at the EcoRI and XhoI sites,  
 respectively."

BASE COUNT 180 a 79 c 108 g 194 t 1 others  
 ORIGIN

Query Match 35.1%; Score 189.8; DB 9; Length 562;  
 Best Local Similarity 68.9%; Pred. No. 2.4e-24;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 AATGCTGCTCTTGTGCTTCATGCGATTTGCTATCTTGGCAAGATGCTTTGTGCC 60  
 DB 27 ATGGCTGCTTCATTTCTTTCATGCGATTTTGTGCTTCGCAATGATGCTTTGTACC 86  
 QY 61 TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAACACACATTTCTGGAATGCG 117  
 DB 87 TATGAGGTGCAAGCTAGAGAAATTTGCAAAAGACCAACCAATTTCCAGATTAATGT 146  
 QY 118 ATTACCAACCAACATGAGAAAGCTTATCAAGTGAATTAATTAAGTGTGTCATGT 177  
 DB 147 TTATGAGCTCATCATGTGAAATATTTGATCAAGAAATTAATTAAGTGTGAGCATTTGT 206

QY 178 AGCAAAATCTCTGAGAGGTGCTTATGATTAAGCCATGTGTGTTGATGAGAAATGACT 237  
 DB 207 AGCAAAATCTCTGAGAGGTGCTTATGATTAAGCCATGTGTAT---TGACAAATCTCA 263  
 QY 238 AAAACGAGACTGAATTTTGGCTGAGAGCAAAATCTTGGCTGACGTTGCTTGA 297  
 DB 264 AGTGAAGTTAAAGCACTTTGGGTGAGAGCAAAATCTTGAATGATGATGCTTGA 323  
 QY 298 GAAGAGATTAATGATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346  
 DB 324 GAAGAGATTAATGATGAGTAAATTAAGTGAAGTAAATTAAGATTTGAGTCAAAA 383  
 QY 347 GATACACATTAATTAAGTTTCTTACCTTTCTTAAAGTAC---TAATGTGTGTTT 403  
 DB 384 AAACAAATTAATTAAGTTTCTTACCTTTCTTAAAGTACCTTGTGTGTGTAG 443  
 QY 404 AATTGCTTTAGTACCTTTTATTAACCTTTAATTAAGTGGCACTTCAT 456  
 DB 444 TATTGGCTTATGATGACCATTTGACACATTAATTAAGTTGTGACACATCAT 496

RESULT 13  
 A1898009 564 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST267452 tomato ovary, TAMU Lycopersicon esculentum cDNA clone  
 DEFINITION CLEDD31M2, mRNA sequence.  
 ACCESSION A1898009  
 VERSION A1898009.1 GI:5603911  
 KEYWORDS EST.  
 SOURCE Lycopersicon esculentum (tomato)  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanales; Solanum; Lycopersicon.  
 1 (bases 1 to 564)  
 Alcala,J., Vredalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,  
 Liang,F., Upton,J., Roming,C.M., Craven,M.B., Fujii,C.Y., Bowman,  
 C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,  
 S.D. and Giovannoni,J.  
 Generation of ESTs from tomato carpel tissue  
 Unpublished  
 CONTACT: CUGI  
 CLEMSON UNIVERSITY Genomics Institute  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1..564  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEDD31M2"  
 /tissue\_type="carpel"  
 /dev\_stage="5 days pre-anthesis to 5 days post-anthesis"  
 /lab\_host="X11-Blue MRF"  
 /clone\_1lb="tomato ovary, TAMU"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; CLEDD - Tomato Carpel EST Library. Oligodt-primed and  
 directionally cloned cDNA in vector lambda ZAP II with 5'  
 and 3' ends located at the EcoRI and XhoI sites,  
 respectively."

BASE COUNT 185 a 80 c 108 g 190 t 1 others  
 ORIGIN

Query Match 35.1%; Score 189.8; DB 9; Length 564;  
 Best Local Similarity 68.9%; Pred. No. 2.4e-24;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 AATGCTGCTCTTGTGCTTCATGCGATTTGCTATCTTGGCAAGATGCTTTGTGCC 60

Db 27 ATGGCTGCTTCATTTTCTTCATGACATTTTGTGCTTGCGCATGATGCTTTTGTACC 86  
 QY 61 TATGAGGTGCAAGCTAGAGAA--TGCAAAACAGAAAGCAACATTTCTGGAATATGC 117  
 Db 87 TATGAGGTGCAAGCTAGAGAAATTTGCAAGACCAAGCCAAATTTCCAGAGATTATGT 146  
 QY 118 ATTACCAAAACCAATGACGAAAAGCTTGATGAGTGAAGAAATTTACTGATGCTATGT 177  
 Db 147 TTTATGACTCATCATGAGAAAATATGTATCAAGAAATTTACTGCTGAGCATGTGT 206  
 QY 178 AGCAAAATCTCAGAGGTGCTGCTATGTAAGCATGCTGCTGTTGATGAGAAAGTACT 237  
 Db 207 AGCAAACTCCAAAGAGATGCTATGCTACCTAGCCATGTAAT--TGACAAATCTCA 263  
 QY 238 AAAACAGAGCTGAAATTTTGGCTGAGAGACAAAACCTTGGCTGACGCTTTGCTTGA 297  
 Db 264 AGTGAAGTTAAAGCACTTTGGTGAGAGACAAAACCTTAAGTGAAGTTGTGCTTGA 323  
 QY 298 GAAGGATTAATGATTAATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346  
 Db 324 GAAGGATTAATGATTAATTAATTAAGATTAGAGAAATTAAGATTTTGAAGTCAAA 383  
 QY 347 GTATCACATTAATTAAGTTCTACTTTCTTAAAGGTAGC--TAATGTTGTGTTT 403  
 Db 384 AAAACAAATTAATTAAGTTGTGCTTTCTTATTAAGGTAGCTTGATGTTGTGTAG 443  
 QY 404 AATTGGCTTTAGTACCTTTTATTAACATTTAATTAAGTGGCACTTCAAT 456  
 Db 444 TATTGGCCTATAGTAGCCATTGACACATTAATTAAGTTGTGTGACACATCAT 496  
 RESULT 14  
 B1931665 712 bp mRNA linear EST 18-OCT-2001  
 LOCUS EST51554 tomato flower, 8 mm to preanthesis buds Lycopersicon  
 DEFINITION esculentum cDNA clone CTC020P19 5' end, mRNA sequence.  
 B1931665  
 ACCESSION B1931665.1 GI:16246137  
 VERSION EST.  
 KEYWORDS  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 712)  
 van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,  
 Utecherack,T., Van Ken,S., Ronting,C.M., Niemman,W., Fraser,C.M.,  
 Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
 Generation of ESTs from tomato flower tissue, buds 8 mm -  
 preanthesis:  
 unpublished  
 CONTACT: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 This clone is available through the Clemson University Genomics  
 Institute  
 Seq primer: T3.  
 Location/Qualifiers  
 1..712  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CTC020P19"  
 /tissue\_type="flower"  
 /dev\_stage="buds 8mm to preanthesis"  
 /clone\_lib="tomato flower, 8 mm to preanthesis buds"  
 /note="Vector: pBluescript SK(-), Site\_1: EcoRI; Site\_2:  
 XhoI; supplier: Cornell University; sequencing: the  
 Institute for Genomic Research; Flower buds and flowers  
 were taken from greenhouse plants (4-8 wks old, TA496).

BASE COUNT 235 a 114 c 125 g 238 t  
 ORIGIN  
 Query Match 35.1%; Score 189.8; DB 12; Length 712;  
 Best Local Similarity 68.9%; Pred. No. 2,2e-24;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;  
 QY 1 ATGGCTGCTTCCTTGCTGCTTCATGAGCATTTGCTATCTTGAGAGATGCTTTGTGCC 60  
 Db 58 ATGGCTGCTTCATTTTCTTATGAGCATTTTGTGCTTGGAATATGCTCTTGTACC 117  
 QY 61 TATGAGGTGCAAGCTAGAGAA--TGCAAAACAGAAAGCAACATTTCTGGAATATGC 117  
 Db 118 TATGAGGTGCAAGCTAGAGAAATTTGCAAGACCAAGCCAAATTTCCAGAGATTATGT 177  
 QY 118 ATTACCAAAACCAATGACGAAAAGCTTGATGAGTGAAGAAATTTACTGATGCTATGT 177  
 Db 178 TTTATGACTCATCATGAGAAAATATGTATCAAGAAATTTACTGCTGAGCATGTGT 206  
 QY 178 AGCAAAATCTCAGAGGTGCTGCTATGTAAGCATGCTGCTGTTGATGAGAAAGTACT 237  
 Db 207 AGCAAACTCCAAAGAGATGCTATGCTACCTAGCCATGTAAT--TGACAAATCTCA 264  
 QY 238 AAAACAGAGCTGAAATTTTGGCTGAGAGACAAAACCTTGGCTGACGCTTTGCTTGA 297  
 Db 264 AGTGAAGTTAAAGCACTTTGGTGAGAGACAAAACCTTAAGTGAAGTTGTGCTTGA 323  
 QY 298 GAAGGATTAATGATTAATTAATTAAGATTAGAGAAATTAAGATTTTGAAGTCAAA 383  
 Db 324 GAAGGATTAATGATTAATTAATTAAGATTAGAGAAATTAAGATTTTGAAGTCAAA 383  
 QY 347 GTATCACATTAATTAAGTTCTACTTTCTTAAAGGTAGC--TAATGTTGTGTTT 403  
 Db 384 AAAACAAATTAATTAAGTTGTGCTTTCTTATTAAGGTAGCTTGATGTTGTGTAG 443  
 QY 404 AATTGGCTTTAGTACCTTTTATTAACATTTAATTAAGTGGCACTTCAAT 456  
 Db 475 TATTGGCCTATAGTAGCCATTGACACATTAATTAAGTTGTGTGACACATCAT 527  
 RESULT 15  
 AM623541 500 bp mRNA linear EST 18-MAY-2001  
 LOCUS EST321486 tomato flower buds 3-8 mm, Cornell University  
 DEFINITION Lycopersicon esculentum cDNA clone CT0B11C9 5', mRNA sequence.  
 AM623541  
 ACCESSION AM623541.1 GI:7336568  
 VERSION EST.  
 KEYWORDS  
 ORGANISM Lycopersicon esculentum (tomato)  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 500)  
 van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang  
 ,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronting,C.M., Niemman,W.,  
 Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.  
 Generation of ESTs from tomato flower tissue, 3-8 mm buds  
 unpublished  
 CONTACT: CUGI  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 Location/Qualifiers  
 1..500  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"

```
/clone="CTOH11C9"  
/issue_type="flower"  
/dev_stage="3-8mm buds"  
/clone_lib="tomato flower buds 3-8 mm, Cornell University"  
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:  
XhoI; supplier: Tankisley; Flower buds and flowers were  
taken from greenhouse plants (4-8 wks old, TA96). They  
were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."
```

BASE COUNT 160 a 77 c 105 g 158 t

Query Match 34.8%; Score 188.2; DB 9; Length 500;  
Best Local Similarity 68.7%; Pred. No. 4.8e-24;  
Matches 307; Conservative 0; Mismatches 128; Indels 12; Gaps 3;

```
QY 1 ATGGCTGCTCTTGTGCTTCTGATGCAATTTGCTATCTTGGCAAGATGCTTTGTTGCC 60  
DB |||||  
QY 34 ATGGCTGCTCTTGTGCTTCTGATGCAATTTGCTATCTTGGCAAGATGCTTTGTTGCC 93  
DB |||||  
QY 61 TATGAGTGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 117  
DB |||||  
QY 94 TCTGAGTGCAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 153  
DB |||||  
QY 118 ATTACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 177  
DB |||||  
QY 154 TTACCGATTTCATGCTATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 213  
DB |||||  
QY 178 AGCAAAATCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 237  
DB |||||  
QY 214 AGCAAACTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 273  
DB |||||  
QY 238 AAAACAGAGCTGAATTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297  
DB |||||  
QY 274 GAAGTAAAC---ACCTTTGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330  
DB |||||  
QY 298 GAGAGATTAAGTAACTAATTA-----GAGATTGAAGAAATTAAAGATGAGATGATC 351  
DB |||||  
QY 331 GAGAGATTAAGTAACTAATTAAGTAACTAATTAAGTAACTAATTAAGTAACTAATTA 390  
DB |||||  
QY 352 ACACATAATAAGTTTACCTTTCTTAAAGTAACTAATTAAGTAACTAATTAAGTAACT 411  
DB |||||  
QY 391 AAATTAATAAGTAACTTTTAAAGTAACTTAACTTAACTTAACTTAACTTAACTTAACT 450  
DB |||||  
QY 412 TTATAGCTTTTATTAACCTTAA 438  
DB |||||  
QY 451 CTTAGTAGCCATTGATACATTAAATA 477  
DB |||||
```

Search completed: January 28, 2004, 09:54:53  
Job time : 1952 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 08:46:51 ; Search time 21 Seconds  
(without alignments)  
21.554 Million cell updates/sec

Title: US-10-072-809B-18

Sequence: 1 MARSICFMFAFIALRMFLVA.....LAERAKTLAALLLEELMDN 105

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued\_Patents\_AA\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	333	60.3	188	6	5175095-3
2	188	34.1	83	4	US-09-442-631-2
3	122.5	22.2	74	4	US-09-442-631-4
4	121	21.9	47	1	US-08-377-687-28
5	121	21.9	47	1	US-08-656-318A-13
6	121	21.9	47	2	US-08-777-192-28
7	121	21.9	47	2	US-08-956-459-13
8	121	21.9	47	3	US-08-971-982-28
9	120	21.7	75	1	US-08-289-458-2
10	120	21.7	75	2	US-08-761-549-2
11	120	21.7	75	3	US-09-127-646-2
12	110	19.9	80	1	US-08-377-687-49
13	110	19.9	80	2	US-08-777-192-49
14	110	19.9	80	3	US-08-971-982-49
15	110	19.9	80	3	US-09-103-489-20
16	108.5	19.7	79	1	US-08-627-706-15
17	108.5	19.7	79	3	US-09-103-489-15
18	108	19.6	80	1	US-08-377-687-59
19	108	19.6	80	2	US-08-777-192-59
20	108	19.6	80	3	US-08-971-982-59
21	107	19.4	74	1	US-08-543-238-5
22	107	19.4	74	1	US-08-420-526-5
23	99	17.9	48	1	US-08-377-687-29
24	99	17.9	48	1	US-08-656-318A-11
25	99	17.9	48	1	US-08-777-192-29
26	99	17.9	48	2	US-08-956-459-11
27	99	17.9	48	3	US-08-971-982-29

28	97	17.6	47	4	US-09-077-951-21	Sequence 21, Appl
29	95	17.2	74	1	US-08-543-238-2	Sequence 2, Appl
30	95	17.2	74	1	US-08-420-526-2	Sequence 2, Appl
31	90.5	16.4	51	4	US-09-077-951-61	Sequence 61, Appl
32	88.5	16.0	51	4	US-09-077-951-25	Sequence 25, Appl
33	88.5	16.0	51	4	US-09-077-951-46	Sequence 46, Appl
34	88.5	16.0	51	4	US-09-077-951-66	Sequence 66, Appl
35	88.5	16.0	51	4	US-09-077-951-76	Sequence 76, Appl
36	88	15.9	47	1	US-08-377-687-30	Sequence 30, Appl
37	88	15.9	47	1	US-08-656-318A-12	Sequence 12, Appl
38	88	15.9	47	2	US-08-777-192-30	Sequence 30, Appl
39	88	15.9	47	2	US-08-956-459-12	Sequence 12, Appl
40	88	15.9	47	3	US-08-971-982-30	Sequence 30, Appl
41	86.5	15.7	51	4	US-09-077-951-26	Sequence 26, Appl
42	86.5	15.7	72	4	US-09-003-198A-17	Sequence 17, Appl
43	85.5	15.5	46	3	US-08-632-511A-5	Sequence 5, Appl
44	85.5	15.5	46	3	US-09-091-590A-11	Sequence 11, Appl
45	85.5	15.5	46	3	US-09-488-200-5	Sequence 5, Appl

#### ALIGNMENTS

```

RESULT 1
5175095-3
; Patent No. 5175095
; APPLICANT: Martineau, Belinda M., Houck, Catherine M.
; TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/554,195
; FILING DATE: 17-JUL-1990
; SEQ ID NO: 3
; LENGTH: 188
5175095-3

Query Match      60.3%; Score 333; DB 6; Length 188;
Best Local Similarity 62.5%; Pred. No. 18-30; Indels 2; Gaps 2;
Matches 65; Conservative 12; Mismatches 25;

QY      1  MARSICFMFAFIALRMFLVAEYQARE-CKTESNTPGICITKPPCRKACISEKFTDGHG 59
Db      14  MARSICFMFAFIALRMFLVAEYQARE-CKTESNTPGICITKPPCRKACISEKFTDGHG 73
QY      60  SKLQKRCCTKPCVDEKKTGTGAETLAEEATLAALIEEIM 103
Db      74  SKLQKRCCTKPCVDEKKTGTGAETLAEEATLAALIEEIM 116

RESULT 2
US-09-442-631-2
; Sequence 2, Application US/09442631
; Patent No. 6300489
; GENERAL INFORMATION:
; APPLICANT: OH, BOUNG-JUN
; APPLICANT: KO, MOON KYUNG
; APPLICANT: SHIN, BYONGCHUL
; APPLICANT: CHUNG, CHANG HO
; TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIPYRAL DEFENSIN AND
; TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE
; FILE REFERENCE: 1942/44
; CURRENT APPLICATION NUMBER: US/09/442,631
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-442-631-2

Query Match      34.1%; Score 188; DB 4; Length 83;

```

TELEPHONE: (202) 861-3075

TELEFAX: (202) 822-0944  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: p322  
US-08-656-318A-13

Query Match 21.9%; Score 121; DB 1; Length 47;  
Best Local Similarity 42.6%; Pred. No. 4,4e-07;  
Matches 20; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 26 RECKTESNTPGICITKPPCKACISEKFTDGHCSKILRRCLCTKPC 72  
DB 1 RHCELSHRFKGPTCRDSNCASVCETERPSGNGCHGFRRCCTKPC 47

RESULT 6  
US-08-777-192-28  
Sequence 28, Application US/0877192  
Patent No. 5824869  
GENERAL INFORMATION:  
APPLICANT: BROEKAERT, WILLEM F.  
APPLICANT: CAMMUE, BRUNO P.A.  
APPLICANT: OSBORN, RUPERT W.  
APPLICANT: REES, SARAH B.  
APPLICANT: TERRAS, FRANKY R.G.  
APPLICANT: VANDERLEYDEN, JOZEF  
TITLE OF INVENTION: BIOCIDAL PROTEINS  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DABRY & CUSHMAN  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/777,192  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/002,480  
FILING DATE: 04-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-777-192-28

Query Match 21.9%; Score 121; DB 2; Length 47;  
Best Local Similarity 42.6%; Pred. No. 4,4e-07;  
Matches 20; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 26 RECKTESNTPGICITKPPCKACISEKFTDGHCSKILRRCLCTKPC 72  
DB 1 RHCELSHRFKGPTCRDSNCASVCETERPSGNGCHGFRRCCTKPC 47

RESULT 7  
US-08-956-459-13  
Sequence 13, Application US/08956459  
Patent No. 5919918  
GENERAL INFORMATION:  
APPLICANT: BROEKAERT, WILLEM F.  
APPLICANT: CAMMUE, BRUNO P.A.  
APPLICANT: OSBORN, RUPERT W.  
APPLICANT: REES, SARAH B.  
TITLE OF INVENTION: ANTIMICROBIAL PROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY MADISON & SUTRO LLP  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,459  
FILING DATE: 22-OCT-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/656,318  
FILING DATE: 12-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB94/02766  
FILING DATE: 19-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9326424.0  
FILING DATE: 24-DEC-1993  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: p322  
US-08-956-459-13

Query Match 21.9%; Score 121; DB 2; Length 47;  
Best Local Similarity 42.6%; Pred. No. 4,4e-07;  
Matches 20; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 26 RECKTESNTPGICITKPPCKACISEKFTDGHCSKILRRCLCTKPC 72  
DB 1 RHCELSHRFKGPTCRDSNCASVCETERPSGNGCHGFRRCCTKPC 47

RESULT 8  
US-08-971-982-28  
Sequence 28, Application US/08971982  
Patent No. 6187904  
GENERAL INFORMATION:  
APPLICANT: BROEKAERT, WILLEM F.  
APPLICANT: CAMMUE, BRUNO P.A.  
APPLICANT: OSBORN, RUPERT W.  
APPLICANT: REES, SARAH B.  
APPLICANT: TERRAS, FRANKY R.G.  
APPLICANT: VANDERLEYDEN, JOZEF  
TITLE OF INVENTION: BIOCIDAL PROTEINS



NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DAREY & CUSHMAN  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/971,982  
FILING DATE: 17-NOV. 6187904-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/002,480  
FILING DATE: 04-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKUBIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-08-971-982-28

Query Match 21.7%; Score 121, DB 3; Length 47;  
Best Local Similarity 42.6%; Pred. No. 4.4e-07;  
Matches 20; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 26 RECKTESNTPGICITKPPCKACISEKFTDGHCKSLIRRLCTKPC 72  
DB 1 RHCSLSHRFKGPGCTRBSNCASVETFRFSGNCHGRRCFCCKPC 47

RESULT 9  
US-08-289-458-2  
Sequence 2, Application US/08289458  
Patent No. 5608144  
GENERAL INFORMATION:  
APPLICANT: BADEN, Catherine S., DUNSMUIR, Pamela,  
APPLICANT: LEE, Kathleen Y.  
TITLE OF INVENTION: PLANT GP2 PROMOTERS AND USES THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/289,458  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.

REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 12176-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-289-458-2

Query Match 21.7%; Score 120, DB 1; Length 75;  
Best Local Similarity 34.3%; Pred. No. 9.7e-07;  
Matches 23; Conservative 12; Mismatches 30; Indels 2; Gaps 1;

QY 8 MAFILAMLFVAIEV--QARECKTESNTPGICITKPPCKACISEKFTDGHCKSLIR 65  
DB 8 VAITFLMVLIVFATDMMAEAKICALTGNFGLCLSRDCGVCRREGFTDGSCTGRLO 67

QY 66 CLCTKPC 72  
DB 68 CFCCKPC 74

RESULT 10  
US-08-761-549-2  
Sequence 2, Application US/08761549  
Patent No. 5981727  
GENERAL INFORMATION:  
APPLICANT: BADEN, Catherine S., DUNSMUIR, Pamela,  
APPLICANT: LEE, Kathleen Y.  
TITLE OF INVENTION: PLANT GP2 PROMOTERS AND USES THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/761,549  
FILING DATE: 06-DEC-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/289,458  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 12176-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-761-549-2

Query Match 21.7%; Score 120, DB 2; Length 75;  
Best Local Similarity 34.3%; Pred. No. 9.7e-07;  
Matches 23; Conservative 12; Mismatches 30; Indels 2; Gaps 1;



TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-777-192-49

Query Match 19.9%; Score 110; DB 2; Length 80;  
Best Local Similarity 32.5%; Pred. No. 1.5e-05;  
Matches 26; Conservative 16; Mismatches 30; Indels 8; Gaps 4;

QY 1 MARSLCFMAFALILARMLFVAYE---VQARE-CKTESNTPPGICITKPPCKKACIS-EKF 54  
DB 1 MAKFASTIALLPALVLPFAAFEAFTWEAKLCERPSTGSGVCGNNACKKNCINLEKA 60  
QY 55 TDGHCSKIL--RRCLCTKPC 72  
DB 61 RHGSCNYVFAHAKCICYFPC 80

RESULT 14  
US-08-971-982-49  
Sequence 49, Application US/08971982  
Patent No. 6187904

GENERAL INFORMATION:  
APPLICANT: BROEKAERT, WILLEM F.  
CAMMUE, BRUNO P.A.  
OSBORN, RUPERT W.

REES, SARAH B.  
TERRAS, FRANKY R.G.  
VANDERLEIDEN, JOZEF

TITLE OF INVENTION: BIOCIDAL PROTEINS  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.

COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/971,982  
FILING DATE: 17-No. 6187904-1997  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/002,480  
FILING DATE: 04-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16, 773  
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944

INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 49:  
US-08-971-982-49

Query Match 19.9%; Score 110; DB 3; Length 80;  
Best Local Similarity 32.5%; Pred. No. 1.5e-05;  
Matches 26; Conservative 16; Mismatches 30; Indels 8; Gaps 4;

QY 1 MARSLCFMAFALILARMLFVAYE---VQARE-CKTESNTPPGICITKPPCKKACIS-EKF 54  
DB 1 MAKFASTIALLPALVLPFAAFEAFTWEAKLCERPSTGSGVCGNNACKKNCINLEKA 60  
QY 55 TDGHCSKIL--RRCLCTKPC 72  
DB 61 RHGSCNYVFAHAKCICYFPC 80

RESULT 15  
US-09-103-489-20  
Sequence 20, Application US/09103489  
Patent No. 6215048

GENERAL INFORMATION:  
APPLICANT: Liang, Jihong  
APPLICANT: Shah, Dilip M.  
APPLICANT: Wu, Yonnie S.

APPLICANT: Rosenberger, Cindy A.  
TITLE OF INVENTION: Antifungal Polypeptide and Methods for  
Controlling Plant Pathogenic Fungi

NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Charles E. Cohen, Monsanto Company, B4F  
STREET: 700 Chesterfield Village Parkway No. 6215048th  
CITY: St. Louis

STATE: Missouri  
COUNTRY: USA  
ZIP: 63198

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/103,489  
FILING DATE: 24-JUN-1998  
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:  
NAME: Cohen, Charles E.  
REGISTRATION NUMBER: 34,565  
REFERENCE/DOCKET NUMBER: 38-21 (10700)A

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-6224  
TELEFAX: (314) 537-6047

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS:

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-103-489-20

Query Match 19.9%; Score 110; DB 3; Length 80;  
Best Local Similarity 32.5%; Pred. No. 1.5e-05;  
Matches 26; Conservative 16; Mismatches 30; Indels 8; Gaps 4;

QY 1 MARSLCFMAFALILARMLFVAYE---VQARE-CKTESNTPPGICITKPPCKKACIS-EKF 54  
DB 1 MAKFASTIALLPALVLPFAAFEAFTWEAKLCERPSTGSGVCGNNACKKNCINLEKA 60  
QY 55 TDGHCSKIL--RRCLCTKPC 72  
DB 61 RHGSCNYVFAHAKCICYFPC 80

Search completed: January 28, 2004, 08:49:14  
Job time : 22 secs

Thu Jan 29 11:57:51 2004

us-10-072-809b-18.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 08:48:46 : Search time 371 Seconds  
(without alignments)

58,510 Million cell updates/sec

Title: US-10-072-809b-18  
Perfect score: 552  
Sequence: 1 MARSLCFMAFALIAFMLFVA.....LAEEAKTIAALAEIEMDN 105

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCIT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCITUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	105	12	US-10-072-809a-18
2	529	98.8	105	12	US-10-072-809a-20
3	508	92.0	105	12	US-10-072-809a-51
4	468.5	84.9	106	12	US-10-072-809a-52
5	431	78.1	80	12	US-10-072-809a-16
6	395	71.6	72	12	US-10-072-809a-14
7	342	62.0	79	12	US-10-072-809a-15
8	333	60.3	105	12	US-10-072-809a-21
9	274	49.6	47	12	US-10-072-809a-8
10	272	49.3	47	12	US-10-072-809a-25
11	185	33.5	47	12	US-10-072-809a-26
12	163	29.5	78	15	US-10-178-213-878
13	160	29.0	75	15	US-10-178-213-290
14	157	28.4	33	12	US-10-072-809a-12
15	155	28.1	78	15	US-10-178-213-431

16	151	27.4	74	15	US-10-178-213-2	Sequence 2, Appli
17	150.5	27.3	77	15	US-10-178-213-407	Sequence 407, App
18	148	26.8	72	15	US-10-178-213-413	Sequence 413, App
19	148	26.6	78	15	US-10-178-213-83	Sequence 83, Appl
20	147	26.6	84	15	US-10-178-213-386	Sequence 386, App
21	145.5	26.4	79	15	US-10-178-213-293	Sequence 293, App
22	145	26.3	47	15	US-10-178-213-279	Sequence 279, App
23	145	26.3	78	15	US-10-178-213-66	Sequence 86, Appl
24	144	26.1	78	12	US-10-072-809a-23	Sequence 23, Appl
25	143	25.9	78	12	US-10-072-809a-22	Sequence 22, Appl
26	141	25.5	77	15	US-10-178-213-443	Sequence 443, App
27	140.5	25.5	79	15	US-10-178-213-308	Sequence 308, App
28	140	25.4	79	15	US-10-178-213-125	Sequence 125, App
29	140	25.4	82	15	US-10-178-213-41	Sequence 41, Appl
30	139.5	25.3	76	15	US-10-178-213-281	Sequence 281, App
31	139	25.2	47	15	US-10-178-213-84	Sequence 84, Appl
32	138	25.0	77	15	US-10-178-213-446	Sequence 446, App
33	138	25.0	77	15	US-10-178-213-445	Sequence 445, App
34	138	25.0	78	15	US-10-178-213-434	Sequence 434, App
35	137.5	24.9	77	12	US-10-072-809a-24	Sequence 24, Appl
36	137.5	24.9	77	15	US-10-178-213-410	Sequence 410, App
37	137	24.8	78	9	US-09-917-340-54	Sequence 54, Appl
38	137	24.8	78	9	US-09-917-340-55	Sequence 55, Appl
39	136	24.6	47	15	US-10-178-213-87	Sequence 87, Appl
40	135.5	24.5	73	15	US-10-178-213-5	Sequence 5, Appl
41	135	24.5	47	15	US-10-178-213-291	Sequence 291, App
42	135	24.5	73	15	US-10-178-213-347	Sequence 347, App
43	134	24.3	47	15	US-10-178-213-3	Sequence 3, Appl
44	134	24.3	64	15	US-10-178-213-122	Sequence 122, App
45	133.5	24.2	77	15	US-10-178-213-455	Sequence 455, App

#### ALIGNMENTS

RESULT 1  
US-10-072-809a-18  
Sequence 18, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267,271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 18  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Nicotiana glauca

Query Match 100.0%, Score 552, DB 12, Length 105,  
Best Local Similarity 100.0%, Pred. No. 2.66-55,  
Matches 105, Conservative 0, Mismatches 0, Indels 0, Gaps 0,  
Db 61  
1 MARSLCFMAFALIAFMLFVAIVQARECKTESNTPPGICITKPPCKACISKFTDGHCS 60  
1 MARSLCFMAFALIAFMLFVAIVQARECKTESNTPPGICITKPPCKACISKFTDGHCS 60  
61 KILRRCLCTKPCVFPDEKMTKGAETIAEEAKTIAALAEIEMDN 105  
61 KILRRCLCTKPCVFPDEKMTKGAETIAEEAKTIAALAEIEMDN 105  
RESULT 2  
US-10-072-809a-20  
Sequence 20, Application US/10072809A  
Publication No. US20030217382A1

GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
TITLE OF INVENTION: therefor  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267,271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 20  
LENGTH: 105  
TYPE: PRT  
ORGANISM: peptide  
US-10-072-809A-20

Query Match 95.8%; Score 529; DB 12; Length 105;  
Best Local Similarity 95.2%; Pred. No. 1.1e-52;  
Matches 100; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MARSICFMAFALILARMFVAYEVOARECKTESNTFFPGICITKPPCKKACISEKFTDGHCS 60  
Db 1 MARSICFMAFALILARMFVAYEVOARECKTESNTFFPGICITKPPCKKACISEKFTDGHCS 60

Qy 61 KILRRCLCTKPCVDEKMTTGAELIAEAKTLAALLBEIEMDN 105  
Db 61 KILRRCLCTKPCVDEKMTTGAELIAEAKTLAALLBEIEMDN 105

RESULT 3  
US-10-072-809A-51  
Sequence 51, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
TITLE OF INVENTION: therefor  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267,271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 51  
LENGTH: 105  
TYPE: PRT  
ORGANISM: peptide  
US-10-072-809A-51

Query Match 92.0%; Score 508; DB 12; Length 105;  
Best Local Similarity 90.5%; Pred. No. 2.7e-50;  
Matches 95; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MARSICFMAFALILARMFVAYEVOARECKTESNTFFPGICITKPPCKKACISEKFTDGHCS 60  
Db 1 MARSICFMAFALILARMFVAYEVOARECKTESNTFFPGICITKPPCKKACISEKFTDGHCS 60

Qy 61 KILRRCLCTKPCVDEKMTTGAELIAEAKTLAALLBEIEMDN 105  
Db 61 KILRRCLCTKPCVDEKMTTGAELIAEAKTLAALLBEIEMDN 105

RESULT 4  
US-10-072-809A-52  
Sequence 52, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
TITLE OF INVENTION: therefor  
FILE REFERENCE: 18-01

CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267,271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 52  
LENGTH: 106  
TYPE: PRT  
ORGANISM: peptide  
US-10-072-809A-52

Query Match 84.9%; Score 468.5; DB 12; Length 106;  
Best Local Similarity 84.0%; Pred. No. 8.8e-46;  
Matches 89; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Qy 1 MARSICFMAFALILARMFVAYEVOARE-CKTESNTFFPGICITKPPCKKACISEKFTDGHCS 59  
Db 1 MARSICFMAFALILARMFVAYEVOARE-CKTESNTFFPGICITKPPCKKACISEKFTDGHCS 60

Qy 60 SKILRRCLCTKPCVDEKMTTGAELIAEAKTLAALLBEIEMDN 105  
Db 61 SKILRRCLCTKPCVDEKMTTGAELIAEAKTLAALLBEIEMDN 106

RESULT 5  
US-10-072-809A-16  
Sequence 16, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
TITLE OF INVENTION: therefor  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267,271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 16  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Nicotiana glauca  
US-10-072-809A-16

Query Match 78.1%; Score 431; DB 12; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 RECKTESNTFFPGICITKPPCKKACISEKFTDGHCSKILRRCLCTKPCVDEKMTTGAELI 85  
Db 1 RECKTESNTFFPGICITKPPCKKACISEKFTDGHCSKILRRCLCTKPCVDEKMTTGAELI 60

Qy 86 LAEAKTLAALLBEIEMDN 105  
Db 61 LAEAKTLAALLBEIEMDN 80

RESULT 6  
US-10-072-809A-14  
Sequence 14, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and  
TITLE OF INVENTION: therefor  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072,809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267,271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61

SOFTWARE: Patentin version 3.0  
SEQ ID NO 14  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Nicotiana glauca  
US-10-072-809A-14

Query Match 71.6%; Score 395; DB 12; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1.3e-37;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLCFMFAIILAMLFVAYEVOARECKTESNTFPGICITKPPCKKACISEKFTDGHG 60  
DB 1 MARSLCFMFAIILAMLFVAYEVOARECKTESNTFPGICITKPPCKKACISEKFTDGHG 60  
QY 61 KILRRCLCTKPC 72  
DB 61 KILRRCLCTKPC 72

RESULT 7  
US-10-072-809A-50  
Sequence 50, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and v  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072, 809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267, 271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 50  
LENGTH: 79  
TYPE: PRT  
ORGANISM: peptide  
US-10-072-809A-50

Query Match 62.0%; Score 342; DB 12; Length 79;  
Best Local Similarity 77.5%; Pred. No. 1.6e-31;  
Matches 69; Conservative 1; Mismatches 9; Indels 10; Gaps 2;

QY 17 LFVAYEVQARECKTESNTFPGICITKPPCKKACISEKFTDGHGSKILRRCLCTKPCVDE 76  
DB 1 LFVAYEVQARECKTESNTFPGICITKPPCKKACISEKFTDGHGSKILRRCLCTKPCVDE 76  
QY 77 KMTKTGAETLAEAKTLAALILEEETMDN 105  
DB 55 ----TGAETLAEAKTLAALILEEETMDN 79

RESULT 8  
US-10-072-809A-21  
Sequence 21, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and v  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072, 809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267, 271  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 21  
LENGTH: 105  
TYPE: PRT  
ORGANISM: peptide

US-10-072-809A-21

Query Match 60.3%; Score 333; DB 12; Length 105;  
Best Local Similarity 62.5%; Pred. No. 2.4e-30;  
Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;

QY 1 MARSLCFMFAIILAMLFVAYEVOARECKTESNTFPGICITKPPCKKACISEKFTDGHG 59  
DB 1 MARSLCFMFAIILAMLFVAYEVOARECKTESNTFPGICITKPPCKKACISEKFTDGHG 59  
QY 60 SKILRRCLCTKPCVDEKMTKTGAETLAEAKTLAALILEEETMDN 103  
DB 61 SKILRRCLCTKPCVDEKMTKTGAETLAEAKTLAALILEEETMDN 103

RESULT 9  
US-10-072-809A-8  
Sequence 8, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and v  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072, 809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267, 271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 8  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Nicotiana glauca  
US-10-072-809A-8

Query Match 49.6%; Score 274; DB 12; Length 47;  
Best Local Similarity 100.0%; Pred. No. 5e-24;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 RECKTESNTFPGICITKPPCKKACISEKFTDGHGSKILRRCLCTKPC 72  
DB 1 RECKTESNTFPGICITKPPCKKACISEKFTDGHGSKILRRCLCTKPC 72

RESULT 10  
US-10-072-809A-25  
Sequence 25, Application US/10072809A  
Publication No. US20030217382A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.  
TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and v  
FILE REFERENCE: 18-01  
CURRENT APPLICATION NUMBER: US/10/072, 809A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: USSN 60/267, 271  
PRIOR FILING DATE: 2001-02-08  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 25  
LENGTH: 47  
TYPE: PRT  
ORGANISM: peptide  
US-10-072-809A-25

Query Match 49.3%; Score 272; DB 12; Length 47;  
Best Local Similarity 97.9%; Pred. No. 8.4e-24;  
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 26 RECKTESNTFPGICITKPPCKKACISEKFTDGHGSKILRRCLCTKPC 72  
DB 1 RECKTESNTFPGICITKPPCKKACISEKFTDGHGSKILRRCLCTKPC 72

```
RESULT 11
US-10-072-809A-26
; Sequence 26, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and u
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072, 809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US2N 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Depside
US-10-072-809A-26

Query Match          33.5%; Score 185; DB 12; Length 47;
Best Local Similarity 66.7%; Pred. No. 7e-14;
Matches 30; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 28 CKTESNTPGICITPCKKACISEKFTDGHCKILRRCCTKPC 72
DB 3 CKASQTFPGICFMDSCCKKICIKETFTGSHCKLQKCKCTKPC 47

RESULT 12
US-10-178-213-278
; Sequence 278, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensein Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 278
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-178-213-278

Query Match          29.5%; Score 163; DB 15; Length 78;
Best Local Similarity 41.0%; Pred. No. 4e-11;
Matches 32; Conservative 11; Mismatches 29; Indels 6; Gaps 1;
```

```
RESULT 13
US-10-178-213-290
; Sequence 290, Application US/10178213
; Publication No. US20030041348A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro Acevedo, Pedro A.
; APPLICANT: Harvell, Leslie
; APPLICANT: Cahoon, Rebecca
; APPLICANT: McCutchen, Billy Fred
; APPLICANT: Lu, Albert
; APPLICANT: Herrmann, Rafael
; APPLICANT: Wong, James
; TITLE OF INVENTION: Defensein Polynucleotides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/246703
; CURRENT APPLICATION NUMBER: US/10/178,213
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: 60/300,152
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/300,241
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 290
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-178-213-290

Query Match          29.0%; Score 160; DB 15; Length 75;
Best Local Similarity 50.7%; Pred. No. 8.5e-11;
Matches 35; Conservative 6; Mismatches 24; Indels 4; Gaps 2;
```

```
QY 7 FMAPFIALMLFVAEYV--QARECKTESNTPGICITPCKKACISEKFTDGHCKIL 63
DB 8 FPAVALLL-LVVAIEVPAQARECETESERNGLCFVSANAGVGNAGFTGKCSGLK 66

QY 64 RRCLCTKPC 72
DB 67 RSCMCTKPC 75

RESULT 14
US-10-072-809A-12
; Sequence 12, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072, 809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US2N 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-10-072-809A-12

Query Match          28.4%; Score 157; DB 12; Length 33;
Best Local Similarity 100.0%; Pred. No. 7.2e-11;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 15  
US-10-178-213-431

/ Sequence 431, Application US/10178213 /  
/ Publication No. US20030041348A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Simmons, Carl R.  
/ APPLICANT: Navarro Acevedo, Pedro A.  
/ APPLICANT: Harvell, Leslie  
/ APPLICANT: Cahoon, Rebecca  
/ APPLICANT: McCutchen, Billy Fred  
/ APPLICANT: Lu, Albert  
/ APPLICANT: Hermann, Rafael  
/ APPLICANT: Wong, James  
/ TITLE OF INVENTION: Defensein Polynucleotides and Methods of  
/ TITLE OF INVENTION: Use  
/ FILE REFERENCE: 35718/246703  
/ CURRENT APPLICATION NUMBER: US/10/178,213  
/ CURRENT FILING DATE: 2002-06-21  
/ PRIOR APPLICATION NUMBER: 60/300,152  
/ PRIOR FILING DATE: 2001-06-22  
/ PRIOR APPLICATION NUMBER: 60/300,241  
/ PRIOR FILING DATE: 2001-06-22  
/ NUMBER OF SEQ ID NOS: 469  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 431  
/ LENGTH: 78  
/ TYPE: PRT  
/ ORGANISM: Nicotiana benthamiana  
US-10-178-213-431

Query Match 28.1%; Score 155; DB 15; Length 78;  
Best Local Similarity 38.5%; Pred. No. 3.3e-10;  
Matches 30; Conservative 14; Mismatches 28; Indels 6; Gaps 1;

QY 1 MARSICFMAFATLARMLFVAYEV-----QARECKTESNTFGICITKPPCKKACISEKF 54  
DB 1 MARSICFMAFATLARMLFVAYEV-----QARECKTESNTFGICITKPPCKKACISEKF 54  
QY 55 TDGHCSTILRRLCTKPC 72  
DB 61 SGDCRCGLRRRCFCCTRPC 78

Search completed: January 28, 2004, 09:00:13  
Job time : 372 secs



OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MTLVANG 23;  
 RA Yun C.-H., Lee M.C., Eun M.Y.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IR36;  
 RA Yun C.-H., Lee J.-H., Park J.-H., Lee G.-R.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. NIPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa niponbare (GA3) genomic DNA, chromosome 3, PAC  
 RT clone: P0043E01."  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U72842; AAB17095.1; -;  
 DR EMBL; AF044059; AAC00503.1; -;  
 DR EMBL; AF000615; BAA05411.1; -;  
 DR HSSP; P20230; 1GPT.  
 DR Gramene; O24225; -;  
 DR InterPro; IPR02118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott1; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
 SQ SEQUENCE 81 AA; 8757 MW; D073BED5AC6BA27 CRC64;  
 Query Match 26.7%; Score 147.5; DB 10; Length 81;  
 Best Local Similarity 37.0%; Pred. No. 2, 1e-10;  
 Matches 27; Conservative 13; Mismatches 26; Indels 7; Gaps 1;

QY 7 FMAPAILARMLFAVEV-----QARECKTESNTFPQICITKPPCKKACISKFTDGRG 59  
 DB 8 FSAMILMTVTLAATGEMGPGVMAEARTCESQSHRFKGPCARKANCASVCTEGFPDGYC 67  
 QY 60 SKILRCLCTKPC 72  
 DB 68 HGVRRRCWCTKPC 80

Search completed: January 28, 2004, 08:48:08  
 Job time : 37 secs

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Protease inhibitor.  
 OS Glycine max (Soybean).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OC NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Essex; TISSUE=Leaf;  
 RA Maitra N., Cushman J.C.;  
 RT "Characterization of a Drought-Induced Soybean cDNA Encoding a Plant  
 RL Plant Physiol. 118:1536-1536 (1998)."  
 DR EMBL; U12150; AAC97524.1; -  
 DR HSSP; P20230; 1GPT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
 SQ SEQUENCE 79 AA; 8772 MW; 24B241BFF67058E3 CRC64;  
 Query Match 27.3%; Score 150.5; DB 10; Length 79;  
 Best Local Similarity 36.7%; Pred. No. 8.7e-11;  
 Matches 29; Conservative 16; Mismatches 27; Indels 7; Gaps 1;  
 QY 1 MARSICFMFAFIALRMFLVAYEV-----QARECKTESNTFPGICTKPPCKRACISEK 53  
 DB 1 MSRSVPLVSTICVILLVATEMGMPTWVAEATCESQSHRFKPCISDTNCGSVCTER 60  
 QY 54 FTDGHCSKILRRCLCTKPC 72  
 DB 61 FTGCHCRGFRRCFCCTKPC 79  
 RESULT 13  
 Q9C947 PRELIMINARY; PRT; 55 AA.  
 AC Q9C947;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Hypothetical 6.0 kDa protein.  
 GN T7P1.20.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDJIN=21018719; PubMed=1130712;  
 RA Theologis A., Ecker J.R., Palm C.U., Federspiel N.A., Kaul S., White O., Alonso J., Alcafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chen H., Cheuk R.F., Chin C.W., Dunn M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Eggu P., Feldhym T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.U., Koo H.L., Kremenetska I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A., Miltiescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Taiton L.J., Tambunga G., Tortum M.J., Town C.D., Utterback I., Van Aken S., Vaysberg M., Vystotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."

RL Nature 408:816-820 (2000).  
 DR EMBL; AC018908; AAG51654.1; -  
 DR HSSP; P20230; 1GPT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 55 AA; 5977 MW; 167C3B51F9B0F52 CRC64;  
 Query Match 27.0%; Score 149; DB 10; Length 55;  
 Best Local Similarity 50.0%; Pred. No. 9.2e-11;  
 Matches 25; Conservative 7; Mismatches 18; Indels 0; Gaps 0;  
 QY 23 VQARECKTESNTFPGICTKPPCKRACISEKFTDGHCSKILRRCLCTKPC 72  
 DB 6 VEARCTETSSNLFNGPCISSNCANVCNHEGFSDDGFRRCCTKPC 55  
 RESULT 14  
 Q9MB66 PRELIMINARY; PRT; 78 AA.  
 AC Q9MB66;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Thionin like protein.  
 GN NT-THIONIN.  
 OS Nicotiana tabacum (Common tobacco).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Takemoto D., Kawakita K.;  
 RT "Molecular cloning of fungal elicitor inducible genes of Tobacco."  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB034956; BAA95697.1; -  
 DR HSSP; P20230; 1GPT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
 SQ SEQUENCE 78 AA; 8577 MW; 1993CAB5C94C8AE51 CRC64;  
 Query Match 27.0%; Score 148; DB 10; Length 78;  
 Best Local Similarity 37.2%; Pred. No. 1.3e-10;  
 Matches 29; Conservative 13; Mismatches 30; Indels 6; Gaps 1;  
 QY 1 MARSICFMFAFIALRMFLVAYEV-----QARECKTESNTFPGICTKPPCKRACISEK 54  
 DB 1 MANSMPFATVILLVATEMGMPTWVAEATCESQSHRFKPCISDTNCGSVCTER 60  
 QY 55 TQGHCSKILRRCLCTKPC 72  
 DB 61 SGGDCRGFRRCFCCTKPC 78  
 RESULT 15  
 Q24225 PRELIMINARY; PRT; 81 AA.  
 AC Q24225;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)  
 DE Proteinase inhibitor.  
 GN RPI OR RGP19.  
 OS Oryza sativa (Rice).

## RESULT 9

```

O9SEMI PRELIMINARY; PRT; 84 AA.
ID O9SEMI
AC O9SEMI
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Gamma-thionin 1 precursor.
GN THION1.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hanbyul; TISSUE=leaf;
RA Hwang B.K., Lee S.C., Kim Y.J., Hong J.K.;
RT "Molecular cloning and pathogen-induced expression of a thionin gene
in pepper plants.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF11869; AAF18936.1;
DR InterPro; IPR002118; Gamma-thionin.
DR Pfam; PF003614; Knott1.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knott; 1.
KW SIGNAL.
FT CHAIN 1 21 POTENTIAL.
FT SIGNAL 22 84 GAMMA-THIONIN 1.
SQ SEQUENCE 84 AA; 9470 MW; 038211731C04E493 CRC64;

Query Match 35.2%; Score 194.5; DB 10; Length 84;
Best Local Similarity 44.4%; Pred. No. 3.7e-16;
Matches 44; Conservative 10; Mismatches 28; Indels 17; Gaps 2;

QY 1 MARSICMAFAILARMLFVAVEVQARECKTSNTPPGICITKPCRKACI-SEKPTDQHC 59
DB 1 MARSIFYMFLVATMTFVAAYGVGGEICKELTPVACSSDPICQKLCMEKEXEDQHC 60
QY 60 SKILRCICTKPCVDEKMTKGAETLAEBAKTAAAL 98
DB 61 FTILSKCLCMKRC-----NAKTLATELL 83

RESULT 10
O9W4V6 PRELIMINARY; PRT; 78 AA.
ID O9W4V6
AC O9W4V6
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Proteinase inhibitor precursor.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hanbyul; TISSUE=leaf;
RA Hwang B.K., Do H.M.;
RT "Differential expression of pepper proteinase inhibitor gene in
response to pathogen infection, abiotic elicitors and osmotic stresses
in Capsicum annuum.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF442388; AAL35366.1;
DR InterPro; IPR002118; Gamma-thionin.
DR Pfam; PF003614; Knott1.
DR ProDom; PD002594; Gamma-thionin; 1.
DR ProDom; PD002594; Gamma-thionin; 1.

```

```

DR SMART; SM00505; Knott; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
KW SIGNAL.
FT SIGNAL 1 20 POTENTIAL.
SQ SEQUENCE 78 AA; 8689 MW; 63BF6DAA79EB4C5D CRC64;

Query Match 29.9%; Score 165; DB 10; Length 78;
Best Local Similarity 39.7%; Pred. No. 1.4e-12;
Matches 31; Conservative 14; Mismatches 27; Indels 6; Gaps 1;

QY 1 MARSICMAFAILARMLFVAVEVQARECKTSNTPPGICITKPCRKACISEKF 54
DB 1 MARSIFYMFLVATMTFVAAYGVGGEICKELTPVACSSDPICQKLCMEKEXEDQHC 60
QY 55 TDGHCCKILRCICTKPC 72
DB 61 SGGDCKRGFRRCCTRPC 78

RESULT 11
O9LEG6 PRELIMINARY; PRT; 76 AA.
ID O9LEG6
AC O9LEG6
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Protease inhibitor II.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldman K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
annotation. 0:0-0(2002).
RL Genome Biol. 0:0-0(2002).
RN (2)
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldman K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085425; AAM62652.1;
DR InterPro; IPR002118; Gamma-thionin.
DR Pfam; PF003614; Knott1.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knott; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
SQ SEQUENCE 76 AA; 8343 MW; E10E218FEAC6E5BE CRC64;

Query Match 28.5%; Score 157.5; DB 10; Length 76;
Best Local Similarity 42.9%; Pred. No. 1.2e-11;
Matches 30; Conservative 11; Mismatches 28; Indels 1; Gaps 1;

QY 3 RSLCMAFAILARMLFVAVEVQARECKTSNTPPGICITKPCRKACISEKPTDQHC 62
DB 8 KSALELLFMTLATVM-GPVREARTCTSSNLFNGPCLSSNCANVCNHEGFSDDGCRGF 66
QY 63 LRRCICTKPC 72
DB 67 RRRCTKRPC 76

RESULT 12
O39807 PRELIMINARY; PRT; 79 AA.
ID O39807
AC O39807
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Lamiales; Solanales; Solanaceae; Petunia.  
 CX NCBI\_TaxID=4102;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Old Glory Blue;  
 RA Lay F.T., Brugliera F., Anderson M.A.;  
 RT "Isolation and properties of floral defensins from Nicotiana glauca and  
 Petunia hybrida";  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF507976; AAN64750.1;  
 SQ SEQUENCE 101 AA; 11049 MW; 855AFE2BD4052D0B CRC64;  
 Query Match 58.1%; Score 320.5; DB 10; Length 101;  
 Best Local Similarity 60.4%; Pred. No. 1.5e-31;  
 Matches 64; Conservative 10; Mismatches 25; Indels 7; Gaps 2;  
 QY 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 58  
 DB 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 60  
 QY 59 CSKILRRCLCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIMD 104  
 DB 61 CSKILRRCLCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIMD 101  
 RESULT 6  
 Q8H6Q1 PRELIMINARY; PRT; 103 AA.  
 ID Q8H6Q1  
 AC Q8H6Q1  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Floral defensin-like protein 1.  
 GN D1.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Lamiales; Solanales; Solanaceae; Petunia.  
 CX NCBI\_TaxID=4102;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Old Glory Blue;  
 RA Lay F.T., Brugliera F., Anderson M.A.;  
 RT "Isolation and properties of floral defensins from Nicotiana glauca and  
 Petunia hybrida";  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF507975; AAN64750.1;  
 SQ SEQUENCE 103 AA; 11361 MW; A8B629A3E06A0D01 CRC64;  
 Query Match 54.8%; Score 302.5; DB 10; Length 103;  
 Best Local Similarity 57.3%; Pred. No. 2.4e-29;  
 Matches 59; Conservative 10; Mismatches 33; Indels 1; Gaps 1;  
 QY 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 60  
 DB 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 60  
 QY 61 KIIRRCCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIM 103  
 DB 61 KIIRRCCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIM 102  
 RESULT 7  
 Q9XHE3 PRELIMINARY; PRT; 107 AA.  
 ID Q9XHE3  
 AC Q9XHE3  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Putative gamma-thionin.  
 OS Capsicum chinense.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Lamiales; Solanales; Solanaceae; Capsicum.  
 CX NCBI\_TaxID=80379;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. habanero;  
 RA MEDLINE=99372458; PubMed=10447467;  
 RA Aluru M., Curry J., O'Connell M.A.;  
 RT "Nucleotide sequence of a defensin or gamma-thionin-like gene  
 (Accession No. AF128239) from habanero chile (PGR 99-070).";  
 RL Plant Physiol. 120:633-633(1999).  
 DR EMBL; AF128239; AAD21200.1;  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott1; 1.  
 SQ SEQUENCE 107 AA; 11964 MW; B173320A754284C4 CRC64;  
 Query Match 51.4%; Score 284; DB 10; Length 107;  
 Best Local Similarity 57.5%; Pred. No. 4.8e-27;  
 Matches 61; Conservative 9; Mismatches 32; Indels 4; Gaps 3;  
 QY 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 57  
 DB 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 60  
 QY 58 HCSKILRRCLCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIM 103  
 DB 61 HCSKILRRCLCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIM 105  
 RESULT 8  
 Q9SEM4 PRELIMINARY; PRT; 84 AA.  
 ID Q9SEM4  
 AC Q9SEM4  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Thionin-like protein.  
 GN PEPTI.  
 OS Capsicum annuum (Bell pepper).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Lamiales; Solanales; Solanaceae; Capsicum.  
 CX NCBI\_TaxID=4072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20064969; PubMed=10598099;  
 RA Oh B.U., Ko M.K., Kostenyuk I., Shin B., Kim K.S.;  
 RT "Coexpression of a defensin gene and a thionin-like gene via different  
 signal transduction pathways in pepper and Coleletotrichum  
 gloeosporioides interactions";  
 RL Plant Mol. Biol. 41:313-319(1999).  
 DR EMBL; AF112443; AAF16413.1;  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott1; 1.  
 SQ SEQUENCE 84 AA; 9484 MW; 8182116639E4F166 CRC64;  
 Query Match 35.8%; Score 197.5; DB 10; Length 84;  
 Best Local Similarity 45.5%; Pred. No. 1.6e-16;  
 Matches 45; Conservative 9; Mismatches 28; Indels 17; Gaps 2;  
 QY 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 59  
 DB 1 MARSICPMFAIILAMLFVAYEVARCKTESNTPGICITKPPCKRACIS--EKTDDG 60  
 QY 60 SKILRRCLCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIM 98  
 DB 61 SKILRRCLCTKPCVDFDEKMTKGAEILAEAKTAAALAEELIM 83

024105 ID 024105 PRELIMINARY; PRT; 105 AA.  
 AC 024105;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Gamma-thionin.  
 GN NETH102.  
 OS Nicotiana excelsior.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Assteridae; lamids; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_Taxid=61185;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RA Yamada S.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA [2]  
 RP SEQUENCE FROM N.A.  
 RA McKendree W.L., Doostdar H., McColium T.G., Mayer R.T.;  
 RT "CDNA cloning and expression of a gene (Accession No. Z97064) from  
 RT Citrus paradisi roots similar to bacterial YRN1 and HEAT10 proteins  
 RT and an mRNA from Brassica oleracea that is wound and dark inducible  
 RT (PGR97-127)."  
 RL Plant Physiol. 115:314-314(1997).  
 DR EMBL; AB005266; BAA21114.1;  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott1; 1.  
 DR PROSITE; PS00940; GAMMA\_THIONIN; 1.  
 SQ SEQUENCE 105 AA; 11635 MW; B7C586CE8DB565DF CRC64;  
 Query Match 92.0%; Score 508; DB 10; Length 105;  
 Best Local Similarity 90.5%; Pred. No. 1.4e-54;  
 Matches 95; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MARSICPMFAIATLAEVAVYEQARECKTESNTFPGICITKPPCRKACISEKTTDGHG 60  
 DB 1 MARSICPMFAIATLAEVAVYEQARECKTESNTFPGICITKPPCRKACISEKTTDGHG 60  
 QY 61 KILRRCLCTKPCVPECKMTKGAETLAEAKTLAALIEERIMDN 105  
 DB 61 KILRRCLCTKPCVPECKMTKGAETLAEAKTLAALIEERIMDN 105  
 RESULT 3  
 024104 ID 024104 PRELIMINARY; PRT; 79 AA.  
 AC 024104;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Gamma-thionin (Fragment).  
 GN NETH101.  
 OS Nicotiana excelsior.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Assteridae; lamids; Solanales; Solanaceae; Nicotiana.  
 OC NCBI\_Taxid=61185;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RA Yamada S., Komori T., Imaseki H.;  
 RT "CDNA cloning of gamma-thionin from Nicotiana excelsior."  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB005265; BAA21113.1;  
 DR HSSP; P20230; 1GPT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott1; 1.

FT NON TER 1 1  
 SQ SEQUENCE 79 AA; 8785 MW; FB507B957D78D285 CRC64;  
 Query Match 62.0%; Score 342; DB 10; Length 79;  
 Best Local Similarity 77.5%; Pred. No. 2.6e-34;  
 Matches 69; Conservative 1; Mismatches 9; Indels 10; Gaps 2;  
 QY 17 LFAVAYEQARECKTESNTFPGICITKPPCRKACISEKTTDGHGKILRRCLCTKPCVPE 76  
 DB 1 LFAVAYEQARECKTESNTFPGICITKPPCRKACISEKTTDGHGKILRRCLCTKPCVPE 76  
 QY 77 KMTKGAETLAEAKTLAALIEERIMDN 105  
 DB 77 KMTKGAETLAEAKTLAALIEERIMDN 105  
 DB 55 ----TGAEETLAEAKTLAALIEERIMDN 79  
 RESULT 4  
 040128 ID 040128 PRELIMINARY; PRT; 105 AA.  
 AC 040128;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Flower-specific gamma-thionin-like protein/acidic protein  
 DE precursor.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamids; Solanales; Solanaceae; Solanum.  
 OC NCBI\_Taxid=4081;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN=VF36; TISSUE=Distil;  
 RC MEDLINE=95375233; PubMed=7647301;  
 RA Milligan S.B., Gasser C.S.;  
 RT "Nature and regulation of pistil-expressed genes in tomato."  
 RL Plant Mol. Biol. 28:691-711(1995).  
 DR EMBL; U20591; AAA80496.1;  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knott1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; Gamma-thionin; 1.  
 DR SMART; SM00505; Knott1; 1.  
 KW Signal.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 73 GAMMA-THIONIN-LIKE PROTEIN.  
 FT CHAIN 74 105 ACIDIC PROTEIN.  
 SQ SEQUENCE 105 AA; 11914 MW; ADG9B7CEB620E814 CRC64;  
 Query Match 60.3%; Score 333; DB 10; Length 105;  
 Best Local Similarity 62.5%; Pred. No. 4.5e-33;  
 Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;  
 QY 1 MARSICPMFAIATLAEVAVYEQARECKTESNTFPGICITKPPCRKACISEKTTDGHG 59  
 DB 1 MARSICPMFAIATLAEVAVYEQARECKTESNTFPGICITKPPCRKACISEKTTDGHG 59  
 QY 60 KILRRCLCTKPCVPECKMTKGAETLAEAKTLAALIEERIMDN 103  
 DB 60 KILRRCLCTKPCVPECKMTKGAETLAEAKTLAALIEERIMDN 103  
 DB 61 SKIQRKCLCTKPCVPECKMTKGAETLAEAKTLAALIEERIMDN 103  
 RESULT 5  
 08H600 ID 08H600 PRELIMINARY; PRT; 101 AA.  
 AC 08H600;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Floral defensin-like protein 2.  
 GN D2.  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 08:46:51 : Search time 35 Seconds  
(without alignments)  
774.158 Million cell updates/sec

Title: US-10-072-809b-18  
Perfect score: 552  
Sequence: 1 MARSICFMAFALIAFMLFVA.....LAEBAKTIAALLEEIMDN 105

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SPTREMBL 23:  
1: sp\_archea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	546	98.9	105	10	Q8GTMO
2	508	92.0	105	10	Q24105
3	342	62.0	79	10	Q24104
4	333	60.3	105	10	Q40128
5	320.5	58.1	101	10	Q8H6Q0
6	302.5	54.8	103	10	Q8H6Q1
7	284	51.4	107	10	Q9XHE3
8	197.5	35.8	84	10	Q9SEH4
9	194.5	35.2	84	10	Q9SEH1
10	165	29.9	78	10	Q8W4V6
11	157.5	28.5	76	10	Q8H6Q6
12	150.5	27.3	79	10	Q38807
13	149	27.0	55	10	Q9C947
14	147.5	26.7	81	10	Q9MB66
15	142.5	25.8	87	10	Q24225
16					Q94873

17	136	24.6	72	10	Q9XG53	Q9XG53 lycopersico
18	136	24.6	78	10	Q945D8	Q945D8 castanea sa
19	135	24.5	81	10	Q948T2	Q948T2 pyrus pyrif
20	134	24.3	77	10	P82788	P82788 arabidopsis
21	131	23.7	87	10	Q948T4	Q948T4 pyrus pyrif
22	129	23.4	49	10	Q93WS9	Q93WS9 musa acumin
23	128.5	23.3	73	10	Q9FUP3	Q9FUP3 phaseolus c
24	125.5	22.7	77	10	Q39403	Q39403 brassica ca
25	122	22.1	73	10	Q9F014	Q9F014 citrus para
26	122	22.1	113	10	Q942U6	Q942U6 cryza sativ
27	119	21.6	82	10	Q8L698	Q8L698 triticum ae
28	118	21.4	77	10	Q8H766	Q8H766 eleais guin
29	118	21.4	83	10	Q8GTU2	Q8GTU2 picea abies
30	117.5	21.3	75	10	Q39894	Q39894 glycine max
31	116	21.0	83	10	Q40779	Q40779 picea abies
32	112	20.3	73	10	Q9FPP8	Q9FPP8 arabidopsis
33	112	20.3	80	10	Q9F122	Q9F122 arabidopsis
34	106	19.2	80	10	Q9F231	Q9F231 arabidopsis
35	106	19.2	80	10	Q9F338	Q9F338 eutrema was
36	104	18.8	80	10	Q941N7	Q941N7 brassica ol
37	103	18.7	42	10	Q9M6P5	Q9M6P5 helianthus
38	103	18.7	80	10	Q8H6K0	Q8H6K0 brassica ra
39	101	18.3	41	10	Q9M7B6	Q9M7B6 helianthus
40	100	18.1	75	10	Q8W434	Q8W434 vigna radia
41	99	17.9	71	5	Q8IRD7	Q8IRD7 droserophila
42	97	17.6	91	10	Q9AY29	Q9AY29 zea mays (m
43	95	17.2	91	10	Q9AY30	Q9AY30 zea mays (m
44	95	17.2	91	10	Q9AY28	Q9AY28 zea mays (m
45	94.5	17.1	82	10	Q39999	Q39999 hordeum vul

## ALIGNMENTS

RESULT 1	Q8GTMO	PRELIMINARY;	PRT;	105 AA.
ID	Q8GTMO			
AC	Q8GTMO:			
DT	01-MAR-2003 (TREMREL 23, Created)			
DT	01-MAR-2003 (TREMREL 23, Last sequence update)			
DT	01-MAR-2003 (TREMREL 23, Last annotation update)			
DE	Flower-specific defensin precursor.			
GN	NAD1.			
OS	Nicotiana glauca (Winged tobacco) (Persian tobacco).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.			
OX	NCBI_TaxID=4087;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RA	Lay P.T., Schirra H.J., Scanlon M.J., Anderson M.A., Craik D.J.;			
RT	"The three-dimensional solution structure of NAD1, a new floral			
RT	defensin from Nicotiana glauca and its application to a homology model			
RT	of the crop defense protein alfap."			
RL	J. Mol. Biol. 0:0-0(2003).			
DR	EMBL; AF509566; AAW0399.1; -			
FT	CHAIN	26	72	FLOWER-SPECIFIC DEFENSIN.
SQ	SEQUENCE	105 AA;	11722 MM;	DAYF41736CEBAC3 CRC64;
Query Match		98.9%;	Score 546;	DB 10;
Best Local Similarity		99.0%;	Pred. No. 3.1e-59;	Length 105;
Matches 104;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;
QY	1 MARSICFMAFALIAFMLFVAEVOARECKTESNTFPGICITKPKKACISEKFTDGHCS 60			
DB	1 MARSICFMAFALIAFMLFVAEVOARECKTESNTFPGICITKPKKACISEKFTDGHCS 60			
QY	61 KIIPRCITKPCVDEDMTKGAEILAEKTLAALLEEIMDN 105			
DB	61 KIIPRCITKPCVDEDMTKGAEILAEKTLAALLEEIMDN 105			
RESULT 2				

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 28, 2004, 08:44:46 ; Search time 219 Seconds

(Without alignments)  
6668.475 Million cell updates/sec

Title: US-10-072-809B-17

Perfect score: 541  
Sequence: 1 atgctcgcctcctctgtctt.....tggttaaaaaaaaaaaaaa 541

Scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N.Geneseq\_19JUn03.\*

```
1: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1880.DAT.*
2: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1881.DAT.*
3: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1882.DAT.*
4: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1883.DAT.*
5: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1884.DAT.*
6: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1885.DAT.*
7: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1886.DAT.*
8: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1887.DAT.*
9: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1888.DAT.*
10: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1889.DAT.*
11: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1890.DAT.*
12: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1891.DAT.*
13: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1892.DAT.*
14: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1893.DAT.*
15: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1894.DAT.*
16: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1895.DAT.*
17: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1896.DAT.*
18: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1897.DAT.*
19: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1898.DAT.*
20: /SIDSI/gcgdata/genseq/genseqn-emb1/NA1899.DAT.*
21: /SIDSI/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/genseq/genseqn-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	490.8	90.7	566	20	AAV70142
2	377.6	69.8	558	20	AAV70143
3	362.8	67.1	456	12	AAV70141
4	189.8	35.1	564	12	AAQ10262
5	189.8	35.1	564	14	AAQ34940
6	189.8	35.1	564	17	AAI31823
7	189.8	35.1	564	18	AAI48813
8	189.8	35.1	564	24	ABZ21970

9	147.4	27.2	3528	17	AAI31824
10	147.4	27.2	3528	18	AAI48816
11	147.4	27.2	3528	24	ABZ21971
12	147.4	27.2	4383	12	AAQ10263
13	147.4	27.2	4383	12	AAQ10319
14	147.4	27.2	4383	14	AAQ35143
15	147.4	27.2	4383	14	AAQ34941
16	95.6	17.7	506	22	AAI38987
17	95.6	17.7	548	24	AAI12800
18	47.4	8.8	19634	25	ABZ10016
19	47.4	8.8	19634	25	ABZ10162
20	45.2	8.4	34548	24	ABL70603
21	44.6	8.2	13326	24	ABL33712
22	43.4	8.0	283	24	ABL75289
23	43.4	8.0	5611	24	ABQ67070
24	43.2	8.0	8056	25	ABZ10246
25	42.4	7.8	1501	25	ABZ10188
26	42.4	7.8	6042	24	ABL33592
27	42.4	7.8	8056	25	ABZ10100
28	42.2	7.8	297	24	ABL75367
29	42	7.8	1692	20	AAI99563
30	42	7.8	64061	24	ABA92787
31	41.8	7.7	6137	24	ABL70128
32	41.8	7.7	6137	24	ABL34451
33	41.8	7.7	6250	24	ABN80214
34	41.8	7.7	13133	24	ABK31230
35	41.8	7.7	18283	24	ABL70502
36	41.8	7.7	18283	24	ABK61363
37	41.6	7.7	506	23	ABV58067
38	41.6	7.7	1338	24	ABZ16790
39	41.4	7.7	6754	24	ABL70346
40	41.4	7.7	6754	24	AAI63305
41	41.4	7.7	17280	22	AAI46772
42	41	7.6	6665	22	AAI46772
43	41	7.6	6665	24	ABL33082
44	41	7.6	6665	24	ABK28129
45	41	7.6	10433	24	ABL32378

#### ALIGNMENTS

RESULT 1					
AAV70142	AAV70142 standard; cDNA to mRNA; 566 BP.				
ID	AAV70142	standard; cDNA to mRNA; 566 BP.			
AC	AAV70142				
XX	XX				
DT	03-FEB-1999	(first entry)			
XX	XX				
DE	Nicotiana excelsior thionine gene NETH12.				
XX	XX				
KW	Nicotiana excelsior; Nicotiana paniculata; thionine; NETH1; NETH2;				
KM	NETH1; salt stress; resistance; ds.				
OS	Nicotiana excelsior.				
XX	XX				
Key	Location/Qualifiers				
FT	33..350				
FT	CDS	/*tag= a			
XX	XX				
PN	JP10295380-A.				
XX	XX				
PD	10-NOV-1998.				
XX	XX				
PF	23-APR-1997;	97JP-0120179.			
XX	XX				
PR	23-APR-1997;	97JP-0120179.			
XX	XX				
PA	(NISR) JAPAN TOBACCO INC.				
XX	XX				
DR	WPI: 1999-038278/04.				
DR	P-PSDB; AAW83132.				

Tomato pZ130 genom  
Calgene Lambda 140  
Plasmid pZ130 cDNA  
pZ130 contg. Calge  
Calgene lambda 140  
Calgene lambda 140  
Calgene lambda 140  
Pepper defensein pr  
DNA encoding Capsi  
Haematopoietic cel  
Haematopoietic cel  
Chemically treated  
Human immune syste  
Corn tassal-derive  
Human angiogenesis  
Haematopoietic cel  
Haematopoietic cel  
Human immune syste  
Haematopoietic cel  
Corn tassal-derive  
Nucleic acid seque  
Buchnera sp. genom  
Chemically treated  
Human metastasis a  
Human chemically m  
Signal transductio  
Chemically treated  
Human gene regulat  
Human prostate exp  
Arabidopsis thalia  
Chemically treated  
Human gene regulat  
Human gene regulat  
Tumour suppressor  
Chemically pretrea  
Human immune syste  
DNA transcription  
Human immune syste

XX Thionine gene derived by salt stress - used to deliver improved salt  
 PT stress to plants  
 XX  
 PS Claim 3; Page 5; 6pp; Japanese.  
 XX  
 CC The present sequence represents a thionine gene from *Nicotiana excelsior*  
 CC derived *NetH12*. The thionine protein has an effect of improving the salt  
 CC stress resistance of a plant. The gene can improve the salt stress  
 CC resistance of a plant.  
 XX  
 SQ Sequence 566 BP; 178 A; 94 C; 108 G; 186 T; 0 other;

Query Match 90.7%; Score 490.8; DB 20; Length 566;  
 Best Local Similarity 94.9%; Pred. No. 1.7e-110;  
 Matches 507; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATGGCTGCTCCCTTGGCTTCATGCAATTTGCTATGCTGGCAAGATGCTCTTTGTCG 60  
 DB 33 ATGGCTGCTCCGCTGCTTCATGCAATTTGCTATGCTGGCAAGATGCTCTTTGTCG 92  
 QY 61 TATGAGTGCAGCTAGAGATGCAAAACAGAAAGCAACATTTCTGGAATATGCAAT 120  
 DB 93 TATGATGTGAGAGCTTAAGATGCAAAACAGAAAGCAATATCCCTGGAATATGCAAT 152  
 QY 121 ACCAAACCAACGCAAAAGCTGTATGATGAGAAATTAATGATGATGCTATGTCAG 180  
 DB 153 ACCAAACCAACGCAAAAGCTGTATGATGAGAAATTAATGATGATGCTATGTCAG 212  
 QY 181 AAATCTCAGAAAGTGCCTATGATGATGATGATGATGATGATGATGATGATGATG 240  
 DB 213 AAATCTCAGAAAGTGCCTATGATGATGATGATGATGATGATGATGATGATGATG 272  
 QY 241 ACAGAGACTGAATTTTGGCTGAGAGCAAAACCTTGGCTGACCTTGGCTGAGAA 300  
 DB 273 ACAGAGACTGAATTTTGGCTGAGAGCAAAACCTTGGCTGACCTTGGCTGAGAA 332  
 QY 301 GAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 360  
 DB 333 GAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 392  
 QY 361 AAAGTTTCTACCTTTCTTAAAGTGTAGCTATGATGATGATGATGATGATGATG 420  
 DB 393 AAAGTTTCTACCTTTCTTAAAGTGTAGCTATGATGATGATGATGATGATGATG 452  
 QY 421 CTTTATATACCTTTAAATAGTGTGACATGCTTGTGCAATCTTGCACTAAGT 480  
 DB 453 CTTTATATACCTTTAAATAGTGTGACATGCTTGTGCAATCTTGCACTAAGT 512  
 QY 481 TTATTTGTGATCTTTAAATAGAAATGACCTTGTATGCTCTTGTGTTAAAAAA 534  
 DB 513 TTATTTGTGATCTTTAAATAGAAATGACCTTGTATGCTCTTGTGTTAAAAAA 566

RESULT 2  
 AAV70143 standard; cDNA to mRNA; 558 BP.  
 ID AAV70143;  
 AC AAV70143;  
 XX  
 DT 03-FEB-1999 (first entry)  
 XX  
 DE *Nicotiana paniculata* thionine gene *NpTH1*.  
 XX  
 KW *Nicotiana excelsior*; *Nicotiana paniculata*; thionine; *NetH1*; *NetH2*;  
 XX *NpTH1*; salt stress; resistance; ds.  
 XX  
 OS *Nicotiana paniculata*.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 48..368  
 FT /+tag= a  
 XX

PN JP10295380-A.  
 XX  
 PD 10-NOV-1998.  
 XX  
 XX  
 PE 23-APR-1997; 97JP-0120179.  
 XX  
 XX 23-APR-1997; 97JP-0120179.  
 XX  
 ER 23-APR-1997; 97JP-0120179.  
 XX  
 PA (NISB) JAPAN TOBACCO INC.  
 XX  
 XX WPI; 1999-038278/04.  
 DR P-PSDB; AAM83133.  
 XX  
 XX  
 PT Thionine gene derived by salt stress - used to deliver improved salt  
 PT stress to plants  
 XX  
 PS Claim 5; Page 5-6; 6pp; Japanese.  
 XX  
 CC The present sequence represents a thionine gene from *Nicotiana*  
 CC *paniculata* derived *NpTH1*. The thionine protein has an effect of  
 CC improving the salt stress resistance of a plant. The gene can improve  
 CC the salt stress resistance of a plant.  
 CC  
 SQ Sequence 558 BP; 175 A; 87 C; 112 G; 184 T; 0 other;

Query Match 69.8%; Score 377.6; DB 20; Length 558;  
 Best Local Similarity 86.2%; Pred. No. 8.1e-83;  
 Matches 464; Conservative 0; Mismatches 44; Indels 30; Gaps 3;

QY 1 ATGGCTGCTCCCTTGGCTTCATGCAATTTGCTATGCTGGCAAGATGCTCTTTGTCG 60  
 DB 48 ATGGCTGCTCCCTTGGCTTCATGCAATTTGCTATGCTGGCAAGATGCTCTTTGTCG 107  
 QY 61 TATGAGTGCAGCTAGAGATGCAAAACAGAAAGCAACATTTCTGGAATATGCAAT 117  
 DB 108 TATGAGTGCAGCTAGAGATGCAAAACAGAAAGCAACATTTCTGGAATATGCAAT 167  
 QY 118 ATTACCAACCAACGCAAAAGCTGTATGATGAGAAATTAATGATGATGCTATGTCAG 177  
 DB 168 ATTACCAACCAACGCAAAAGCTGTATGATGAGAAATTAATGATGATGCTATGTCAG 227  
 QY 178 AGCAAAATCCCTCAGAAAGTGCCTATGATGATGATGATGATGATGATGATGATG 237  
 DB 228 AGCAAAATCCCTCAGAAAGTGCCTATGATGATGATGATGATGATGATGATGATG 287  
 QY 238 AAAACAGAGCTGAATTTTGGCTGAGAGCAAAACCTTGGCTGACCTTGGCTGAA 297  
 DB 288 AAAACAGAGCTGAATTTTGGCTGAGAGCAAAACCTTGGCTGACCTTGGCTGAA 347  
 QY 298 GAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 357  
 DB 348 GAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 407  
 QY 358 AATAAGTTTCTACCTTTCTTAAAGTGTAGCTATGATGATGATGATGATGATGATG 417  
 DB 408 AATAAG-TGCTGCTTTCTTAAAGTGTAGCTATGATGATGATGATGATGATGATG 466  
 QY 418 AGCTTTTATACCTTTAAATAGTGTGACATGCTTGTGCAATCTTGCACTAAGT 477  
 DB 467 AGCTTTTATACCTTTAAATAGTGTGACATGCTTGTGCAATCTTGCACTAAGT 512  
 QY 478 AGTTATTTGTGATCTTTAAATAGAAATGACCTTGTATGCTCTTGTGTTAAAAAA 535  
 DB 513 -----TACTTTAATGAAATGATCTTGTATGCTCTTGTGTTAAAAAA 558

RESULT 3  
 AAV70141 standard; cDNA to mRNA; 456 BP.  
 ID AAV70141;  
 AC AAV70141;  
 XX  
 DT 03-FEB-1999 (first entry)  
 XX



[illegible]

QY	529	AAAAAA	534
Db	451	AAAAAA	456
RESULT 4:			
ID	AAQ10262		
XX	AAQ10262 standard; cDNA, 564 BP.		
XX	AAQ10262;		
XX	25-MAR-2003 (updated)		
DT	04-APR-1991 (first entry)		
DE	Ovary tissue transcriptional factor DNA clone pz130.		
XX	Ovary tissue transcriptional factor; DNA construct; probe;		
KM	clone pz130; ds.		
XX			
OS	Lycopersicon esculentum UC82B.		
XX			
FT	Key	Location/Qualifiers	
FT	msc_feature	447..564	
FT		/*tag= a	
XX		/label= pz7_probe	
XX	EP409629-A.		
XX	23-JAN-1991.		
XX			
PF	19-JUL-1990; 90EF-0307925.		
XX			
XX	19-JUL-1989; 89US-0382518.		
XX			
PA	(CALJ ) CALGENE INC.		
PI	Martineau B, Houck CM;		
XX			
XX	WPI; 1991-024191/04.		
DR	P-PSDB; AAR10310.		
XX			
PT	New ovary tissue transcriptional factors - modify transcription		
PT	in tomato plant ovaries for use as mol. probes		
XX			
PS	Disclosure; Fig 1: 21bp; English.		
XX			
CC	The tomato-derived transcriptional initiation region which regulates		
CC	the expression of the sequence corresp. to the pz130 clone is		
CC	considered ovary-specific. Sequences hybridisable to the pz130		
CC	clone, e.g. probe pz7, show abundant mRNA, esp. at the early stages		
CC	of anesthesia. The message is expressed in ovary integument and ovary		
CC	outer pericary tissue and is not expressed, or at least not readily		
CC	detectable, in other tissues or at any other stage of fruit		
CC	development. The native function of the amino acid sequence		
CC	encoded by the structural gene comprising pz130 is unknown.		
CC	See also AAQ10263-64.		
CC	(Updated on 25-MAR-2003 to correct PA field.)		
SQ	Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;		
QY	Query Match	35.1%; Score 189.8; DB 12; Length 564;	
Db	Best Local Similarity	68.9%; Pred. No. 6.5e-37;	
	Matches 32;	Conservative 0; Mismatches 127; Indels 20; Gaps 4	
QY	1	ATGGCTGCGCTCTTGTGCTTCATGGCAATTTCGTCATGTCATGCTCTTTGTTGCC	60
Db	40	ATGGCTGCGCTTCATTTCTTTCATGGCAATTTCGTCATGTCATGTCATGCTCTTTGTTACC	99
QY	61	TATGAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACATTTCTGGAATATGC	117
Db	100	TATGAGGTGCAAGCTCAGCAATTTGCAAAACCAAGCAAACTTTCCAGGATTATGT	159
QY	118	ATTACCAACCCACATGCAGAAAGCTGTATCATGATGAAATTACTATGTCATTTGT	177

```

Db      160 TTTATGACTCATCATGTAAGAAAATTTGATCAAGAGAAATTTCTGTGACATTGT 219
QY      178 AGCAAAATCTCTGAAAGTGCCCTATGACTTAAGCCATGTGTGTTGATGAGAAGTACT 237
Db      220 AGCAAACTCCAAAGGAAGTGTATGACTTAAGCCATGTGTATTT---TGACAAATCTCA 276
QY      238 AAAACGAGCTGAATTTTGGCTGAGGAGCAAAATCTGGCTGACCTTGTCTGAA 297
Db      277 AGTGAAGTTAAGAACTTGGGTGAGGAGCAAAATCTTAAGTAAGTTGTGTTGAA 336
QY      298 GAGAGATAATGATTAATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346
Db      337 GAGAGATTATGATGAGTATTAATTAAGAGTTAAATTAAGATTGAGTGCATAA 396
QY      347 GTATCAACATTAATTAAGTTTCACTTTCTTAAAGTGTAGC---TAATGTGCTTTT 403
Db      397 AAAACAAATTAATTAAGTGTGCTTTCTTATTAAGGTAAGCTTGTGATGTGTAG 456
QY      404 AATTGCTTTTAAAGCTTTTATTAACCTTTAAATTAAGTGCACCTTCAAT 456
Db      457 TATTGGCCTATGAGGCACTTGCACATTAATTAAGTTTGTGACATCATTT 509

```

## RESULT 5

AAQ34940 ID AAQ34940 standard; DNA; 564 BP.

AC AAQ34940;

DT 25-MAR-2003 (updated)  
DT 19-MAY-1993 (first entry)

DE p2130 coding sequence.

XX cDNA: clone p2130; anthesis; tomato; ovary; integument; outer pericarp;  
KM fruit; development; transcription; initiation; region; modulation;  
KM ovary-specific; fruit product; exogenous; phenotype; ds.

OS Lycopersicon esculentum.

XX Key Location/Qualifiers  
FT misc\_RNA 447..564  
FT /tag= a  
FT /note= "Corresponds to p27 cDNA".

PN US5175095-A.

XX 29-DEC-1992.

PF 17-JUL-1990; 90US-0554195.

PR 19-JUL-1989; 89US-0382518.

PR 17-JUL-1990; 90US-0554195.

PA (CALJ) CALGENE INC.

PI Houck CM, Martineau BW;

DR WPI; 1993-026940/03.

DR P-PSDB; AAR30779.

PT DNA constructs contg. tomato p2130 transcriptional initiation

PT region - useful for modulation of endogenous fruit prodn. and for

PT prodn. of exogenous prods.

XX Disclosure; Fig 1; 18pp; English.

XX The sequence given shows the DNA sequence of cDNA clone p2130. This

CC sequence is expressed during the early stages of anthesis in tomatoes.

CC The message is expressed in ovary integument and ovary outer pericarp

CC tissue. It is not readily detectable in other tissues or at other

CC stages of fruit development. The transcription initiation region

CC associated with this gene is therefore considered to be ovary-  
CC specific. The actual function of the p2130 polypeptide is unknown.  
CC The transcription initiation region can be used for modulation of  
CC endogenous fruit products, for production of exogenous products and  
CC for modification of the phenotype of fruit and fruit products.  
XX (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;

Query Match 35.1%; Score 189.8; DB 14; Length 564;

Best Local Similarity 68.9%; Pred. No. 6; Se-37; Indels 20; Caps 4;

Matches 326; Conservative 0; Mismatches 127;

QY 1 ATGGCTGCTCTTGTGCTTCATGAGCAATTTGATATCTGCAAGATGCTTTGTTGCC 60

Db 40 ATGGCTGCTTCATTTCTTCATGAGCAATTTGATATCTGCAAGATGCTTTGTTGCC 99

QY 61 TATGAGTGCAGCTAGAGAA---TGCAAAACGAAAGCAACATTTCTGGAATATGC 117

Db 100 TATGAGTGCAGCTAGAGAAATTTGCAAGCAACCAACTTCCAGATTAATGT 159

QY 118 ATTACCAACCAACATGCAAGAAAGCTTGTATCAGTGAATTTACTGATGTCATTGT 177

Db 160 TTTATGACTCATCATGTAAGAAAATTTGATCAAGAGAAATTTCTGTGACATTGT 219

QY 178 AGCAAAATCTCTGAAAGTGCCCTATGACTTAAGCCATGTGTGTTGATGAGAAGTACT 237

Db 220 AGCAAACTCCAAAGGAAGTGTATGACTTAAGCCATGTGTATTT---TGACAAATCTCA 276

QY 238 AAAACAGAGCTGAATTTTGGCTGAGGAGCAAAATCTGGCTGACCTTGTCTGAA 297

Db 277 AGTGAAGTTAAGAACTTTGGGTGAGGAGCAAAATCTTAAGTAAGTTGTGTTGAA 336

QY 298 GAGAGATAATGATTAATTAATTAAGATTAGAGAAATTAAGAT-----GCA 346

Db 337 GAGAGATTATGATGAGTATTAATTAAGAGTTAAATTAAGATTGAGTGCATAA 396

QY 347 GTATCAACATTAATTAAGTTTCACTTTCTTAAAGTGTAGC---TAATGTGCTTTT 403

Db 397 AAAACAAATTAATTAAGTGTGCTTTCTTATTAAGGTAAGCTTGTGATGTGTAG 456

QY 404 AATTGCTTTTAAAGCTTTTATTAACCTTTAAATTAAGTGCACCTTCAAT 456

Db 457 TATTGGCCTATGAGGCACTTGCACATTAATTAAGTTTGTGACATCATTT 509

## RESULT 6

AAT31823 ID AAT31823 standard; cDNA; 564 BP.

AC AAT31823;

DT 25-MAR-2003 (updated)

DT 14-SEP-1996 (first entry)

XX Tomato p2130 cDNA clone.

XX Ovary; ovule; fruit; tomato; cotton; melanin; vector;

XX transgenic plant; ss.

OS Lycopersicon esculentum cv. UC82B.

XX Key Location/Qualifiers

FT CDS 1..357

FT /tag= a

PN US5530185-A.

XX 25-JUN-1996.

XX 29-DEC-1992; 92US-0998158.

XX 29-DEC-1992; 92US-0998158.

XX 29-DEC-1992; 92US-0998158.

PR 19-JUL-1989; 89US-0382518.  
 PR 17-JUL-1990; 90US-0554195.  
 XX  
 PA (CALJ) CALGENE INC.  
 XX  
 PI Martineau BM, Reilley AA, Stalker DM;  
 XX  
 DR WPI, 1996-308822/31.  
 DR P-PSDB; AAR97559.  
 XX  
 PT DNA construct for expressing melanin synthesis gene in plant ovule  
 PT cells - contains promoter from the tomato p2130 gene, also binary  
 PT vector and transgenic plants, esp. cotton, contg. construct  
 XX  
 PS Example 3; Fig 1A-B; 25pp; English.  
 XX  
 CC The tomato p2130 clone contains a 564 bp insert of cDNA (AAT31823)  
 CC detected only in a tomato cDNA library. Prepd. from pre-anthesis  
 CC RNA. It was isolated by screening a library prepd. from cDNA  
 CC of pre-anthesis stage ovaries with probes made from pre-anthesis  
 CC mRNA, leaf mRNA and young seedling mRNA. The insert was used to  
 CC isolate the corresponding genomic clone (AAT31824). The p2130  
 CC transcriptional initiation region is considered to be ovary-specific.  
 CC It can be utilised in DNA constructs for the expression of  
 CC heterologous genes, partic. in early fruit development, and esp. for  
 CC expression of a melanin synthesis gene in transgenic cotton. The  
 CC native function of the p2130 gene product (AAR97559) is unknown.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 SQ Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;  
 SQ

Query Match 35.1%; Score 189.8; DB 17; Length 564;  
 Best Local Similarity 68.9%; Pred. No. 6.5e-37;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 ATGCTGCTCTCTGCTTCATGCAATTTGCTATCTTGCGAAGAGTCTTTGTTGCC 60  
 XX  
 DB 40 ATGCTGCTCTCTCTGCTTCATGCAATTTGCTATCTTGCGAAGAGTCTTTGTTACC 99  
 XX  
 QY 61 TATAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACATTTCTCGAATATGC 117  
 XX  
 DB 100 TATAGGTGAGAGCTCAGCAAAATTTGCAAGCAACCACTTTCCAGATATATGT 159  
 XX  
 QY 118 ATTACCAACCAACATGAGAAAGCTTGATCATGAGAAATTTACTGATGTCATGT 177  
 XX  
 DB 160 TTTATGAGCTCATGTGAGAAATTTATGATCAAGAAATTTACTGTCGACATGT 219  
 XX  
 QY 178 AGCAAAATCTCAGAGGTGCTTATGTAAGCATGTGTGTTGATGAGAAATGACT 237  
 XX  
 DB 220 AGCAAACTCCAAAGAGAGTGTCTATGCACTAAGCATGTGTAT---TGCAAAATCTCA 276  
 XX  
 QY 238 AAAACAGAGCTGAAATTTGGCTGAGAGCAAAAATTTGGCTGACGTTTGCTTGA 297  
 XX  
 DB 277 AGTGAAGTTAAGCACTTTGGTGAGAGCAAAAATTTAGTGAAGTTGTGCTTGA 336  
 XX  
 QY 298 GAAAGATATGATTAATACTAATTAAGATTGAAGAAATTAAGAT-----GCA 346  
 XX  
 DB 337 GAAAGATATGATGAGATTAATACTAATTAAGAGTTAAATTAAGATTTGAGTCAAA 396  
 XX  
 QY 347 GTATCACACATTAATTAAGTTTCTACCTTTCTTAAGGTAGC---TAACTGTGTTT 433  
 XX  
 DB 397 AAAACAAATTAATAAAGTTGCTTTCTTATTAAGGTAGCTTGAGATGTTGTAG 456  
 XX  
 QY 404 AATTGGCTTTAGTAGCTTTTATTAACATTTAATAAGTGTGACCTTCAT 456  
 XX  
 DB 457 TATTGGCTTAATAGTACATTTGACATTAATAATTAAGTTGTGACATCAT 509  
 XX

XX  
 DT 14-MAR-1997 (first entry)  
 XX  
 DE cDNA clone p2130 capable of directing ovary-tissue transcription.  
 XX  
 KW Ovary; fruit; colour; pigmentation; cotton; tomato; probe;  
 XX  
 KW promoter; p2130; p27; transgenic plant; ds.  
 XX  
 OS Lycopersicon esculentum cv. UC82B.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 1..357  
 FT /\*tag= a  
 FT misc\_RNA 447..564  
 FT /\*tag= b  
 FT /note= "bases 447-564 correspond to clone p27"  
 XX  
 PN WO9640951-A2.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 07-JUN-1996; 96WO-US09911.  
 XX  
 PR 07-JUN-1995; 95US-0487087.  
 XX  
 PA (CALJ) CALGENE INC.  
 XX  
 PI McBride K, Stalker DM;  
 XX  
 DR WPI; 1997-052341/05.  
 DR P-PSDB; AAW08364.  
 XX  
 PT DNA construct capable of directing ovary-tissue transcription in  
 PT plants - useful for modifying colour phenotype, in e.g. cotton  
 XX  
 PS Example 1; Fig 1; 75pp; English.  
 XX  
 CC A cDNA clone (AAT48813), designated p2130, comprises a tomato  
 CC transcriptional initiation region that is capable of directing  
 CC transcription of a gene of interest specifically in plant ovary  
 CC tissue, partic. early in fruit development, e.g. to modify colour  
 CC phenotype. It can also be used as a molecular probe. To obtain  
 CC p2130, a tomato pre-anthesis stage cDNA library was screened by  
 CC differential hybridisation. Clones p27 and p28 that hybridised  
 CC only to pre-anthesis probes were used to screen a second cDNA  
 CC library, yielding p2130 and p270 (see also AAT48814). A genomic  
 CC clone (AAT48816) was also isolated using p2130 as probe.  
 CC  
 SQ Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;  
 SQ

Query Match 35.1%; Score 189.8; DB 18; Length 564;  
 Best Local Similarity 68.9%; Pred. No. 6.5e-37;  
 Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 ATGCTGCTCTCTGCTTCATGCAATTTGCTATCTTGCGAAGAGTCTTTGTTGCC 60  
 XX  
 DB 40 ATGCTGCTCTCTCTGCTTCATGCAATTTGCTATCTTGCGAAGAGTCTTTGTTACC 99  
 XX  
 QY 61 TATAGGTGCAAGCTAGAGAA---TGCAAAACAGAAAGCAACATTTCTCGAATATGC 117  
 XX  
 DB 100 TATAGGTGAGAGCTCAGCAAAATTTGCAAGCAACCACTTTCCAGATATATGT 159  
 XX  
 QY 118 ATTACCAACCAACATGAGAAAGCTTGATCATGAGAAATTTACTGATGTCATGT 177  
 XX  
 DB 160 TTTATGAGCTCATGTGAGAAATTTATGATCAAGAAATTTACTGTCGACATGT 219  
 XX  
 QY 178 AGCAAAATCTCAGAGGTGCTTATGTAAGCATGTGTGTTGATGAGAAATGACT 237  
 XX  
 DB 220 AGCAAACTCCAAAGAGAGTGTCTATGCACTAAGCATGTGTAT---TGCAAAATCTCA 276  
 XX  
 QY 238 AAAACAGAGCTGAAATTTGGCTGAGAGCAAAAATTTGGCTGACGTTTGCTTGA 297  
 XX  
 DB 277 AGTGAAGTTAAGCACTTTGGTGAGAGCAAAAATTTAGTGAAGTTGTGCTTGA 336  
 XX

```

QY 298 GAAGAGTATGATTAAGATTAGAGATTAGAGAAATTAGAT-----GCA 346
Db 337 GAAGAGTTATGATGAGTATATTAAGAGCTTAAATAGATTGAGTCAAAA 396
QY 347 GTATCACATTAATAAGTTCTTACCTTCTTAAAGTAGC---TAATGTTGTTT 403
Db 397 AAACAAATTAATAAGTGTGCTTCTTATAGGAGTAGTGTGATGTTGTTAG 456
QY 404 AATTGCTTTTATGACCTTTTATACCTTTAAATAAGTGGCACTTCAAT 456
Db 457 TATTGGCTTATATGAGCATTTGACACATTAAATAAGTTTGTGACATCAT 509

RESULT 8
ABZ21970
ID ABZ21970 standard; cDNA; 564 BP.
AC ABZ21970;
XX 28-MAR-2003 (first entry)
XX Plasmid pz130 cDNA sequence pz7.
XX Transcriptional factor; ovary tissue; gene; ss.
XX Unspecified.
XX Key Location/Qualifiers
XX CDS 1..564
XX /tag= a
XX /partial
XX /product= "pz130 protein sequence"
XX /trans_except= (pos:355..357,aa:Xaa)
XX /trans_except= (pos:358..360,aa:Xaa)
XX /trans_except= (pos:376..378,aa:Xaa)
XX /trans_except= (pos:385..387,aa:Xaa)
XX /trans_except= (pos:436..438,aa:Xaa)
XX /trans_except= (pos:454..456,aa:Xaa)
XX /trans_except= (pos:478..480,aa:Xaa)
XX /trans_except= (pos:490..492,aa:Xaa)
XX /trans_except= (pos:541..543,aa:Xaa)
XX /note= "all Xaa's are encoded by stop codons"

CN189856-A.
XX PN
XX 05-AUG-1998.
XX PF 07-JUN-1996; 96CN-0195170.
XX PR 07-JUN-1995; 95US-0480087.
XX PA (CALJ) CALGENE INC.
XX DR WPI, 2002-73392/80.
XX DR P-PSDB; ABP56255.
XX PT Use of transcriptional factors of ovary tissue -
XX PS Example 3; Fig 1A-B; 56pp; Chinese.
XX CC The present invention describes the use of transcriptional factors of
XX CC ovary tissue. The present sequence represents a nucleotide sequence
XX CC which is used in an example from the present invention.
XX SQ Sequence 564 BP; 188 A; 80 C; 107 G; 189 T; 0 other;

Query Match 35.1%; Score 189.8; DB 24; Length 564;
Best Local Similarity 68.9%; Pred. No. 6; se 37;
Matches 326; Conservative 0; Mismatches 127; Indels 20; Gaps 4;

QY 1 ATGGCTGCTCTTGTGCTTATGAGCATTTGCTATCTTGCAAGATGCTTGTGTC 60
|||||

```

```

Db 40 ATGGCTGCTTCCATTTTCTTATGCACTTTTGTGCTTGGCAATGCTCTTTGTTACC 99
QY 61 TATAGGTGCAAGCTAGAGAA---TGCAAAACGAAAGCAACATTTCTGGAATATGC 117
Db 100 TATAGGTAGAGAGCTCAGCAAAATTTGCAAGACCAAGCAAACTTTCCAGGATTAATG 159
QY 118 ATTACCAACCAACATGACAGAAAGCTGTATAGTAGAGAAATTTCTGATGTCATTTG 177
Db 160 TTTATGACATCATATATGAGAAATATTGTATCAAGAGAAATTTACTGTGTGACATTTG 219
QY 178 AGCAAAATCCTCAGAGAGTCTTATGTACTTAACCATGTGTGTTGATGAGAAATGACT 237
Db 220 AGCAAACTCCAAAGGAAGTGTATGCACTTAACCATGTGTATTT---TGCAAAATCTCA 276
QY 238 AAACAGAGCTGAAATTTTGGCTGGAGAGCAAAACTTTGGCTGAGCTTGTGA 297
Db 277 AGTGAAGTTAAGCACTTTGGGTGAGAGCAAAACTCTAAGTGAAGTTGTGCTGAA 336
QY 298 GAAGAGATTAATGATTAATTAATTAAGATTAGAGAAATTAGAT-----GCA 346
Db 337 GAAGAGATTATGATGAGTATATTAATTAAGAGTTAAATTAAGATTTGAGTCAAAA 396
QY 347 GTATCACATTAATAAGTTCTTACCTTCTTAAAGTAGC---TAATGTTGTTT 403
Db 397 AAACAAATTAATAAGTGTGCTTCTTATAGGAGTAGTGTGATGTTGTTAG 456
QY 404 AATTGCTTTTATGACCTTTTATACCTTTAAATAAGTGGCACTTCAAT 456
Db 457 TATTGGCTTATATGAGCATTTGACACATTAAATAAGTTTGTGACATCAT 509

RESULT 9
AAT31824
ID AAT31824 standard; cDNA; 3528 BP.
AC AAT31824;
XX 25-MAR-2003 (updated)
XX DT 14-SEP-1996 (first entry)
XX Tomato pz130 genomic clone Calgene Lambda 140.
XX DE Ovary; ovule; fruit; tomato; cotton; melanin; vector;
XX KW transgenic plant; ss.
XX LY Lycopersicon esculentum cv. UC82B.
XX OS
XX Key Location/Qualifiers
XX FH misc_difference 586
XX FT /tag= a
XX FT /note= "base n at position 586 is unidentified"
XX FT misc_difference 590
XX FT /tag= b
XX FT /note= "base n at position 590 is not identified"
XX FT misc_difference 591
XX FT /tag= c
XX FT /note= "base n at position 591 is not identified"
XX FT misc_difference 663
XX FT /tag= d
XX FT /note= "base n at position 663 is not identified"
XX FT misc_difference 667
XX FT /tag= e
XX FT /note= "base n at position 667 is not identified"
XX FT misc_difference 769
XX FT /tag= f
XX FT /note= "base n at position 769 is not identified"
XX FT misc_difference 2483
XX FT /tag= g
XX FT /note= "base 2483 is given as s in the
XX FT /tag= h
XX FT /note= "base 2485 is given as s in the
XX FT misc_difference 2485
XX FT /tag= h
XX FT /note= "base 2485 is given as s in the
XX FT

```



Db 3062 TGACCTAACGCATGCTGATANT---TCACAAATCTCAAGTACAGTTAAAGCAACTTGGGT 3118

QY 262 GAGAAAGCAAAACCTTGGCTGCGAGCTTGGCTTGAAGAAGAGATATGGATACATAATTA 321

Db 3119 GAGGAAAGCAAAACCTTAAAGTAAAGTGTGCTTTAAGAAAGAGATTATGATGAGNDATTA 3178

QY 322 GAGATTAGAAGAATTTAAGAT-----GCAGTATCACATATATTAAGTTCTA 370

Db 3179 TTAAAGAGAGTTAAATTAAGGATTTTGAAGTGTCAAAAAACAAATTAATTAAGAGTTCG 3238

QY 371 CCTTTCTTAAAGTGAAGC---TAATGTGTGTTTAAATGAGCTTTAGAGCCTTTAT 427

Db 3239 CTTTCTTATTTAGGGAGCTGTGTATGTGTGTTAAGATTTGGCCTATAGAGCATTGCA 3299

QY 428 TACACTTAAATTAAGTGGGCACTTCAAT 456

Db 3299 CACATTAAATTAAGTTGTGACACATCATTT 3327

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	27.2%	Score 147.4	DB 24	Length 3528
Best Local Similarity	67.1%	Pred No. 2,36-26			
Matches 261	Conservative 0	Mismatches 111	Indels 17	Gaps 3	
QY	82	TGCAAAACGAAAGCAACATTTCTGSAATGATCATTCACAAACCCATGACAGAAA	141		
DB	2942	TGCAAAAGCACAAAGCCAACTTTCCACGAGATTATGTTTATGACATCATGTTAGAAA	3001		
QY	142	GCTTATATACGTGAGAAATTTATCTGATGATTTGAGCAAAATCCACAAAGTCCCA	201		
DB	3002	TATTGTATCAAAAGAAATTTACTGCTGACATTTGACCAACTCCAAAGAAAGTCTCA	3061		
QY	202	TGTACTAAGCCATGTGTGTTGATGAGAGATGACTAAAACAGAGCTGAAATTTTGCT	261		
DB	3062	TGACCTAAGCAGATGTATTTT---TACAAAATCTCAAGTAAAGTTAAAGCACTTTGGT	3111		
QY	262	GAGAGACGAAAACTTTGGCTGACGCTTGCTTGAAGAAGAGATATGATTAATCTAATTA	321		

Db	3119	GAGAAACCAAAAACTCAAGAGAAAGTGTGCTTGAAAGAGATTATGATGAGTAATAAA	3178
Oy	322	GAGATTGAGAAATTAAGAT-----GCATATGCACATTAATAAGTTCTA	370
Db	3179	TTAAGTAGGTTAATTAAGATTTTGAGTGTCAAAAAACAAATTAATTAAGTGTGC	3258
Oy	371	CCTTTCTTAAAGTGNCC--TAATGTGTGTTTAATTGGCTTTAGTAGCCTTTTAT	427
Db	3239	CTTTCTTATATAGGAGTGTGATGATTGTGTAGTATGGCCCTATAGTACCATTTGA	3258
Oy	428	TACACTTAATAATAGTGGCAGCTTCAT	456
Db	3239	CACATTAATATAGTTGTGACACATCATT	3327

```
RESULT 12
AAQ10263
ID   AAQ10263 standard; DNA; 4383 BP.
```

DT	25-MAR-2003	(updated)
DT	04-APR-1991	(first entry)
XX	pZ130 contg. <i>Caigense</i> lambda 140 genomic clone.	
DE		
XX	Ovary tissue transcriptional factor; DNA construct; probe,	
KW	clone pZ130; ss.	
XX		
OS	<i>Lycopersicon esculentum</i> UC82B.	
XX		
FH	Key	Location/Qualifiers
FT	intron	2702..2921
FT		/*tag= a
FT	misc_RNA	2599..3382
FT		/*tag= b
FT		//label= pZ130
FT		//note= "overlap"
FT	misc_feature	2567..2567
FT		/*tag= c
FT		//label= transcription_start

PN EP409629-A.  
XX  
XX 23-JAN-1991.  
PD  
XX  
XX 19-JUL-1990; 90EP-0307926.  
PF  
XX 19-JUL-1989; 89US-0382518.  
PR  
XX  
XX  
XX (CALJ ) CALGENE INC.  
PA  
XX  
XX Martineau, B, Houck CM,  
PI  
XX  
XX WFI; 1991-024191/04.  
DR

PT New ovary tissue transcriptional factors - modify transcription  
 PI in tomato plant ovaries for use as mol. probes  
 PR  
 PS Disclosure; Fig 2; 21pp; English.  
 XX  
 XX To prepare pCGN2901, Calgene Lambda 140 was digested with SalI  
 CC and the resulting fragment which contains the pZV-hybridising  
 CC region was inserted into pCGN2015, at the unique SalI site.  
 CC pCGN2902 contains the other SalI fragment (non-pZV-hybridising)  
 CC of the pZ130 genome derived from SalI digestion of  
 CC Calgene Lambda 140 also put into a pCGN2015 construct.  
 CC Plasmid DNA isolated from pCGN2901 was used in the construction  
 CC of a pZ130 Promoter Cassette. The promoter is ovary-specific  
 CC and wound-inducible in leaf cells.  
 CC See also AAO10262-64.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 4383 BP; 1539 A; 620 C; 610 G; 1606 T; 8 other;  
 SQ

Query Match 27.2%; Score 147.4; DB 12; Length 4383;  
 Best Local Similarity 67.1%; Pred. No. 2.5e-26;  
 Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

XX 82 TGCAGAAAGCAAGACACATTTCTGGAATATGATTTACCAACGACATGCGAGAAA 141  
 DB 2942 TGCAGAAAGCAAGACACATTTCTGGAATATGATTTACCAACGACATGCGAGAAA 3001  
 QY 142 GCTTGATGAGGAGAAATTTACGATGCTGATGATGACAAATCTCAGAGGTGCTTA 201  
 DB 3002 TATTGATCAAGAGAAATTTACGATGCTGATGATGACAAATCTCAGAGGTGCTTA 3061  
 QY 202 TGTACTAGGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 261  
 DB 3062 TGCATTAAGCCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3118  
 QY 262 GAGGAGCAAGAACTTTGGCTGACCTTTGCTTGAAGAGATTAATGATTAATTA 321  
 DB 3119 GAGGAGCAAGAACTTTGGCTGACCTTTGCTTGAAGAGATTAATGATTAATTA 3178  
 QY 322 GAGATTAAGAAATTAAGAT-----GAGATCAACATTAATTAAGTTCTA 370  
 DB 3179 TTAAGTGAAGTAAATTAAGATTTGAGTCAAAAAACAAATTAATTAAGTTGCTC 3238  
 QY 371 CCTTTCTTAAGAGTGTAGC---TAATGTTGTTTAAATGCTTTAGTACCTTTTAT 427  
 DB 3239 CTTTCTTAATTAAGAGTGTAGC---TAATGTTGTTTAAATGCTTTAGTACCTTTTAT 3298  
 QY 428 TACACTTAATTAAGTGTGACCTTCAAT 456  
 DB 3299 CACATTAAATTAAGTGTGACCTTCAAT 3327

RESULT 13  
 ID AAQ10319 standard; cDNA; 4383 BP.  
 XX AAQ10319;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 04-APR-1991 (first entry)  
 XX  
 DE Calgene lambda 140 genomic clone.  
 XX  
 KW p2130; cytokinin; ds.  
 XX  
 OS Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH 2567..3382  
 FT misc\_RNA  
 FT /note= "Sequence homologous to p2130 transcript"  
 FT /tag= a  
 FT 2599..3382  
 FT /note= b  
 FT /note= "Sequence homologous to p2130 CDS"  
 FT 2702..2921  
 FT /tag= c  
 FT /note= "Sequence homologous to p2130 intron"  
 XX  
 XX BP409628-A.  
 XX  
 XX 23-JAN-1991.  
 XX  
 XX 19-JUL-1990; 90EP-0307925.  
 XX  
 XX 19-JUL-1989; 89US-0382802.  
 XX  
 XX (CALJ ) CALGENE INC.  
 XX  
 XX Houck CM, Pear JR, Martineau B, Hiatt W;

XX WPI; 1991-024190/04.  
 DR  
 XX  
 PT Modulating endogenous cytokinin levels - regulatory regions are  
 FT transformed into plant cells e.g. fruit to modify pheno-type  
 XX  
 XX PS Disclosure; Fig 3; 39pp; English.  
 XX  
 CC The sequence encodes an enzyme in the cytokinin biosynthetic pathway,  
 CC derived from Z130, and useful in modulating a transformed plant's  
 CC phenotype eg. fruit maturation, ripening etc.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 4383 BP; 1539 A; 620 C; 610 G; 1606 T; 8 other;  
 SQ

Query Match 27.2%; Score 147.4; DB 12; Length 4383;  
 Best Local Similarity 67.1%; Pred. No. 2.5e-26;  
 Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;

XX 82 TGCAGAAAGCAAGACACATTTCTGGAATATGATTTACCAACGACATGCGAGAAA 141  
 DB 2942 TGCAGAAAGCAAGACACATTTCTGGAATATGATTTACCAACGACATGCGAGAAA 3001  
 QY 142 GCTTGATGAGGAGAAATTTACGATGCTGATGATGATGATGATGATGATGATGATGAT 201  
 DB 3002 TATTGATCAAGAGAAATTTACGATGCTGATGATGATGATGATGATGATGATGATGAT 3061  
 QY 202 TGTACTAGGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 261  
 DB 3062 TGCATTAAGCCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3118  
 QY 262 GAGGAGCAAGAACTTTGGCTGACCTTTGCTTGAAGAGATTAATGATTAATTA 321  
 DB 3119 GAGGAGCAAGAACTTTGGCTGACCTTTGCTTGAAGAGATTAATGATTAATTA 3178  
 QY 322 GAGATTAAGAAATTAAGAT-----GAGATCAACATTAATTAAGTTCTA 370  
 DB 3179 TTAAGTGAAGTAAATTAAGATTTGAGTCAAAAAACAAATTAATTAAGTTGCTC 3238  
 QY 371 CCTTTCTTAAGAGTGTAGC---TAATGTTGTTTAAATGCTTTAGTACCTTTTAT 427  
 DB 3239 CTTTCTTAATTAAGAGTGTAGC---TAATGTTGTTTAAATGCTTTAGTACCTTTTAT 3298  
 QY 428 TACACTTAATTAAGTGTGACCTTCAAT 456  
 DB 3299 CACATTAAATTAAGTGTGACCTTCAAT 3327

RESULT 14  
 ID AAQ35143 standard; cDNA; 4383 BP.  
 XX AAQ35143;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 10-MAR-2003 (updated)  
 DT 24-MAY-1993 (first entry)  
 XX  
 DE Calgene lambda 140/p2130 DNA including p2130 gene.  
 XX  
 KW Calgene lambda 140; p2130; expression; cytokinin; plant; ss.  
 XX  
 OS Lycopersicon esculentum.  
 OS  
 XX Bacteriophage lambda.  
 OS  
 XX Key Location/Qualifiers  
 FH 2702..2921  
 FT intron  
 FT /tag= a  
 FT 2598..3382  
 FT /tag= b  
 FT /note= "Overlap between lambda 140 genomic clone  
 FT with p2130 cDNA clone"  
 FT  
 FT misc\_RNA 2567

```

FT      /*tag= c
XX      /note= "start of pz130 gene transcript"
XX
XX      US5177307-A.
XX
XX      05-JAN-1993.
XX
XX      17-JUL-1990; 90US-0554196.
XX
XX      26-MAY-1987; 87US-0054369.
XX      15-MAR-1989; 88US-0168190.
XX      28-APR-1988; 88US-0188361.
XX      19-JUL-1989; 89US-0382802.
XX      17-JUL-1990; 90US-0554196.
XX
XX      (CALJ ) CALGENE INC.
XX
XX      Hatt W, Houck CM, Martineau BM, Pear JR;
XX
XX      WPI, 1993-035826/04.
XX
XX      New plant expression cassettes - contg. DNA encoding an enzyme in
XX      a cytokinin metabolic pathway to modify cytokinin levels in
XX      plants
XX
XX      Disclosure; Fig 3; 29pp; English.
XX
XX      The sequence shows the complete sequence of the region of Calgene
XX      lambda 140 genomic clone that overlaps with the pz130 cDNA clone and
XX      a partial sequence of regions 5' and 3' to that region. It is
XX      possible that small gaps in the sequence may exist upstream from the
XX      Salt site bound at position 808. The clone is used in the
XX      construction of a novel expression cassette comprising in the 5'-3'
XX      direction of transcription, as operably joined components: a
XX      transcriptional and translational initiation region from pz130, a DNA
XX      sequence encoding an enzyme in a cytokinin metabolic pathway and a
XX      transcription termination region. The enzyme is pref. DNA
XX      transferase. The expression cassette can be used to modify a plant
XX      phenotype by modifying cytokinin levels during a specific time and/or
XX      in specific tissues during the development of a plant. Control of
XX      cytokinin concns. can be used to increase the total solids content of
XX      ripe fruit, improve retention of early fruit and alter the time of
XX      onset of ripening. See also AAQ35144-5.
XX      (Updated on 10-MAR-2003 to add missing OS field.)
XX      (Updated on 25-MAR-2003 to correct PF field.)
XX
XX      Sequence 4383 BP; 1540 A; 628 C; 606 G; 1600 T; 9 other;
XX
XX      Query Match 27.2%; Score 147.4; DB 14; Length 4383;
XX      Best Local Similarity 67.1%; Pred. No. 2.5e-26;
XX      Matches 261; Conservative 0; Mismatches 111; Indels 17; Gaps 3;
XX
XX      82 TGCAAAACAGAAAGCAACACATTTCTCGAATAGCATTCACAAACACACATGACAGAAA 141
XX      2942 TGCAAAACAGCAACCACTTCCAGGATTTGTTTATGACATCATCATGTACAAA 3001
XX
XX      142 GCTTGATCAGTGAATTTACTGATGTCATGTAGCAAAATCTCTGAGAGTGCCTA 201
XX      3002 TATGTATCAAGAAATTTACTGCGGACATTTGCAAACTCCAAAGAAAGTGTCTA 3061
XX
XX      202 TGATCTAGCAGCTGTGTTGATGAGAAATGACTTAAACAGAGCTGAAATTTGGCT 261
XX      3062 TGCACTAGCCATGTGTATT--TGACAAATCTCAAGTGAAGTTAAAGCACTTTGGGT 3118
XX
XX      262 GAGGAGCAAAAACCTTTGGCTGACGCTTCTGAGAGAGATTAATGATTAATTA 321
XX      3119 GAGGAGCAAAAACCTTAAGTGAAGTGTGCTTGAAGAAGATTAATGATGAGTAATA 3178
XX
XX      322 GAGATTGAGAAATTTAAGAT-----GCAGTATCAACATTAATTAAGTTTCA 370
XX      3179 TTAAGTAGGTTAATTAAGATTTTGAAGTCAAAAAAACAATTAATTAAGTGTGC 3238
XX      371 CCTTTCTTAAAGTGTAGC--TAATGTGTGTTTAAATGCTTTAGTAGCCTTTTAT 427

```

```

DB      3239 CTTTCTTATAGGCTTGCTGAGTGTGTGATGATGAGCTTATGAGCCATTGA 3298
XX
XX      428 TACCTTAAATAGTGTGACCTTCAAT 456
XX
XX      3299 CACATTAATTAAGTTTGTGAACATCAAT 3327
XX
XX
XX      RESULT 15
XX      ID AAQ34941
XX      AAQ34941 standard; DNA; 4383 BP.
XX
XX      AC AAQ34941;
XX
XX      DT 25-MAR-2003 (updated)
XX      DT 19-MAY-1993 (first entry)
XX
XX      DE Calgene Lambda 140 genomic clone.
XX
XX      KW cDNA; clone pz130; anthesis; tomato; ovary; integument; outer pericarp;
XX      fruit; development; transcription; initiation; region; modulation;
XX      ovary-specific; endogenous; fruit product; exogenous; phenotype;
XX      Lycopersicon esculentum; ss.
XX
XX      OS Synthetic.
XX
XX      FH Key
XX      FT misc_RNA
XX      Location/Qualifiers
XX      /tag= a
XX      /note= "Interrupted pz130 cDNA clone"
XX      intron 2702..2921
XX      /tag= b
XX      FT misc_signal
XX      2567
XX      /tag= c
XX      /function= Start of pz130 gene transcript
XX
XX      US5175095-A.
XX      29-DEC-1992.
XX
XX      PF 17-JUL-1990; 90US-0554195.
XX
XX      PR 19-JUL-1989; 89US-0382518.
XX      17-JUL-1990; 90US-0554195.
XX
XX      PA (CALJ ) CALGENE INC.
XX
XX      PI Houck CM, Martineau BM;
XX
XX      WPI, 1993-026940/03.
XX
XX      DNA constructs contg. tomato pz130 transcriptional initiation
XX      region - useful for modulation of endogenous fruit prods. and for
XX      prodn. of exogenous prods.
XX
XX      Disclosure; Fig 2; 18pp; English.
XX
XX      The sequence given shows the sequence of the region of the Calgene
XX      lambda 140 genomic clone which overlaps with the pz130 cDNA clone and
XX      a partial sequence 5' and 3' to that region. The pz130 sequence is
XX      expressed during the early stages of anthesis in tomato. The message
XX      is expressed in ovary integument and ovary outer pericarp tissue. It
XX      is not readily detectable in other tissues or at other stages of fruit
XX      development. The transcription initiation region associated with
XX      this gene is therefore considered to be ovary-specific. The actual
XX      function of the pz130 polypeptide is unknown. The transcription
XX      initiation region can be used for modulation of endogenous fruit
XX      products, for production of exogenous products and for modification
XX      of the phenotype of fruit and fruit products.
XX      (Updated on 25-MAR-2003 to correct PF field.)
XX
XX      Sequence 4383 BP; 1539 A; 620 C; 610 G; 1606 T; 8 other;
XX

```



Thu Jan 29 11:57:49 2004

us-10-072-809b-17.rng

Page 11

Query Match	27.2%	Score 147.4;	DB:14;	Length 4383;
Best Local Similarity	67.1%;	Pred. No. 2.5e-26;		
Matches 261;	Conservative 0;	Mismatches 111;	Indels 17;	Gaps 3;

QY	82	GGCAAAAAGAGAAAGCAACAATTCCTGGAAATATGCATTACCAAGCAGCAACCTGCAGAAA	141
Db	2942	TGCAGAGACCAAGCCAAACCTTCCAGAAATATATGTTTATGGACTCACTCATGAGAAA	3001
QY	142	GCTTGATCAGTGAGAAAATTTACTGATGCTCATGTGACAAATCTCAGAGGTGCTTA	201
Db	3002	TATGTATCAGAAGAAAATTTACTGCTGACATTTGACCAATCTCAAGAGAGTGCTA	3061
QY	202	TGTACTAAGCCATGTGTGTTGATGAGAAGATGACTAAAACAGAGCTGAAATTTGGCT	261
Db	3062	TGCATTAAGCCATGTGATT---TGACAAAATCTCAAGTGAAGTTAAAGCAACTTGGGT	3118
QY	262	GAGGAAGCAAAAACCTTGCTGCACCTTGCTGGAAGAGATATAGATTAACATAATTA	321
Db	3119	GAGGAAGCAAAAACCTTAAGTAAGTGTGCTTGAGAGAGATATATATGAGATTAATA	3178
QY	322	GAGATTAGAAGAAATTAAGAT-----GCATGATCACATATATAAGTTTCTA	370
Db	3179	TTAAGTAGGTTAAATTAAGATTTTGAATGTCAAAAAAAACAAATTAATAAGTGTTGC	3238
QY	371	CCTTTCTTAAAGTGTAGC---TAATGTGTGTTTAATTGGCTTTTAGTGCTTTAT	427
Db	3239	CTTTTCTTATTAGGTAGCTGTGTGATGTGTGTTAGTATGGCCATATGATGACATTGCA	3298
QY	428	TACACTTAADTAAGTGTGCACTTCAAT	456
Db	3299	CACATTAATTAAGTTGTGACACATCAATT	3327

Search completed: January 28, 2004, 09:22:02  
Job time : 224 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: January 28, 2004, 08:46:51 / Search time 41 Seconds

(without alignments)  
406.495 Million cell updates/sec

Title: US-10-072-809B-18

Perfect score: 552  
Sequence: 1 MARSLCFNAFAILARMLEVA.....LAERAKTLAALLLEETIMDN 105

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Geneseq\_19Jun03.\*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	508	92.0	105 20	AAW83132 Nicotiana excelsior
2	468.5	84.9	106 20	AAW83133 Nicotiana paniculata
3	342	62.0	79 20	AAW83131 Nicotiana excelsior
4	333	60.3	118 17	AAW97559 Tomato PZ130 CDNA-
5	333	60.3	118 18	AAW08354 Tomato ovary-speci
6	333	60.3	179 12	AAW10310 Ovary tissue trans
7	333	60.3	188 14	AAW30779 PZ130 polypeptide
8	333	60.3	23	AAW56255 Plasmid pZ130 prot
9	197.5	35.8	84 22	AAW85079 Pepper defensin pr

10	190.5	34.5	84 23	ABU08325 Capsicum annuum L.
11	159.5	28.9	100 21	AG07868 Arabidopsis thailia
12	159.5	28.9	128 21	AG11734 Arabidopsis thailia
13	157.5	28.5	76 21	AG25022 Arabidopsis thailia
14	149	27.0	55 21	AG07870 Arabidopsis thailia
15	149	27.0	61 21	AG07869 Arabidopsis thailia
16	149	27.0	61 21	AG25023 Arabidopsis thailia
17	143.5	26.0	73 21	AG07998 Arabidopsis thailia
18	143.5	26.0	73 21	AG34807 Arabidopsis thailia
19	140.5	25.5	77 21	AG09775 Arabidopsis thailia
20	138.5	25.1	120 21	ABW24880 Arabidopsis thailia
21	138	25.0	58 21	AG08000 Arabidopsis thailia
22	138	25.0	58 21	AG34809 Arabidopsis thailia
23	138	25.0	61 21	AG07999 Arabidopsis thailia
24	138	25.0	61 21	AG34808 Arabidopsis thailia
25	137.5	24.9	77 21	ABW2460 Arabidopsis thailia
26	137.5	24.9	77 21	ABW24882 Arabidopsis thailia
27	137.5	24.9	77 21	AG04820 Arabidopsis thailia
28	137.5	24.9	77 21	AG38973 Arabidopsis thailia
29	137.5	24.9	101 21	ABW24659 Arabidopsis thailia
30	137.5	24.9	109 21	ABW24881 Arabidopsis thailia
31	137	24.8	78 23	AAU91018 Arabidopsis thailia
32	137	24.8	78 23	AAU91019 Arabidopsis thailia
33	132.5	24.0	64 21	AG04822 Arabidopsis thailia
34	132.5	24.0	64 21	AG38975 Arabidopsis thailia
35	132.5	24.0	73 21	AG04821 Arabidopsis thailia
36	132.5	24.0	73 21	AG38974 Arabidopsis thailia
37	124	22.5	75 22	ABW85080 Arabidopsis thailia
38	121.5	22.0	62 21	ABW24661 Arabidopsis thailia
39	121.5	22.0	62 21	AAW09776 Arabidopsis thailia
40	121	21.9	47 14	AAW3765 Potato gene P132
41	120	21.7	75 17	AAW92752 Pepper group 2 pro
42	120	21.7	75 22	AAW1676 Antifungal protein
43	110	19.9	80 15	AAW57325 Antimicrobial Rs-A
44	110	19.9	80 15	AAW57325 Arabidopsis thailia
45	110	19.9	80 18	AAW19280 Raphanus sativus a

## ALIGNMENTS

RESULT 1  
AAW83132  
ID AAW83132 standard; Protein; 105 AA.  
XX  
AC AAW83132;  
XX  
DT 03-FEB-1999 (first entry)  
XX  
DE Nicotiana excelsior thionine protein NetH12.  
XX  
XX Nicotiana excelsior; Nicotiana paniculata; thionine; NetH1; NetH2;  
KM NetH1; salt stress; resistance.  
XX  
OS Nicotiana excelsior.  
XX  
XX JF10295380-A.  
PN  
XX  
PD 10-NOV-1998.  
XX  
PF 23-APR-1997; 97JP-0120179.  
XX  
PR 23-APR-1997; 97JP-0120179.  
XX  
PA (NISB) JAPAN TOBACCO INC.  
XX  
DR WPI; 1999-038278/04.  
DR N-PSDB; AAW70142.  
XX  
PT Thionine gene derived by salt stress - used to deliver improved salt  
stress to plants  
XX  
PS Claim 3; Page 5; 6pp; Japanese.

XX The present sequence represents a thionine protein from Nicotiana  
 CC excelsior derived Neth12. The thionine protein has an effect of  
 CC improving the salt stress resistance of a plant. The gene can improve  
 CC the salt stress resistance of a plant.  
 XX  
 SQ Sequence 105 AA;

Query Match 92.0%; Score 508; DB 20; Length 105;  
 Best Local Similarity 90.5%; Pred. No. 1,5e-45;  
 Matches 95; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 MARSICMPAFAIILAMLFVAIVEVQARE-CKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 DB 1 MARSICMPAFAIILAMLFVAIVEVQARE-CKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 QY 61 KILRRCLCTKPCVDEKMTKGAEILAEAKTILAAALLEEIMDN 105  
 DB 61 KILRRCLCTKPCVDEKMTKGAEILAEAKTILAAALLEEIMDN 105

## RESULT 2

ID AAW83133 standard; Protein; 106 AA.

AC AAW83133;

DT 03-FEB-1999 (first entry)

DE Nicotiana paniculata thionine protein NpTH11.

KM Nicotiana excelsior; Nicotiana paniculata; thionine; NETH11; NETH12;

KW NpTH11; salt stress; resistance.

OS Nicotiana paniculata.

PN JP10295380-A.

PD 10-NOV-1998.

PF 23-APR-1997; 97JP-0120179.

PR 23-APR-1997; 97JP-0120179.

PA (N1SB ) JAPAN TOBACCO INC.

DR WPI; 1999-038278/04.

DR N-PSDB; AAV70143.

PT Thionine gene derived by salt stress - used to deliver improved salt

PT stress to plants

PS Claim 5; Page 5-6; 6pp; Japanese.

CC The present sequence represents a thionine protein from Nicotiana

CC paniculata derived NpTH11. The thionine protein has an effect of

CC improving the salt stress resistance of a plant. The gene can improve

CC the salt stress resistance of a plant.

XX  
 SQ Sequence 106 AA;

Query Match 84.9%; Score 468.5; DB 20; Length 106;  
 Best Local Similarity 84.0%; Pred. No. 2e-41;  
 Matches 89; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 MARSICMPAFAIILAMLFVAIVEVQARE-CKTESNTFPGICITKPPCKACISEKFTDGHCS 59  
 DB 1 MARSICMPAFAIILAMLFVAIVEVQARE-CKTESNTFPGICITKPPCKACISEKFTDGHCS 60  
 QY 60 SKILRRCLCTKPCVDEKMTKGAEILAEAKTILAAALLEEIMDN 105  
 DB 61 SKILRRCLCTKPCVDEKMTKGAEILAEAKTILAAALLEEIMDN 106

## RESULT 3

ID AAW83131 standard; Protein; 79 AA.

AC AAW83131;

DT 03-FEB-1999 (first entry)

DE Nicotiana excelsior thionine protein Neth11.

KM Nicotiana excelsior; Nicotiana paniculata; thionine; NETH11; NETH12;

KW NpTH11; salt stress; resistance.

OS Nicotiana excelsior.

PN JP10295380-A.

PD 10-NOV-1998.

PF 23-APR-1997; 97JP-0120179.

PR 23-APR-1997; 97JP-0120179.

PA (N1SB ) JAPAN TOBACCO INC.

DR WPI; 1999-038278/04.

DR N-PSDB; AAV70141.

PT Thionine gene derived by salt stress - used to deliver improved salt

PT stress to plants

PS Claim 1; Page 4; 6pp; Japanese.

CC The present sequence represents a thionine protein from Nicotiana

CC excelsior derived Neth11. The thionine protein has an effect of

CC improving the salt stress resistance of a plant. The gene can improve

CC the salt stress resistance of a plant.

XX  
 SQ Sequence 79 AA;

Query Match 62.0%; Score 342; DB 20; Length 79;

Best Local Similarity 77.5%; Pred. No. 2,4e-28;

Matches 69; Conservative 1; Mismatches 9; Indels 10; Gaps 2;

QY 17 LFVAIVEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCSKILRRCLCTKPCVDE 76  
 DB 1 LFVAIVEVQARECKTESNTFPGICITKPPCKACISEKFTDGHCSKILRRCLCTKPCVDE 76  
 QY 77 KMTKTGAELAEAKTILAAALLEEIMDN 105  
 DB 55 ----TGAETLAEAKTILAAALLEEIMDN 79

## RESULT 4

ID AAR97559 standard; Protein; 118 AA.

AC AAR97559;

DT 25-MAR-2003 (updated)

DT 14-SEP-1996 (first entry)

DE Tomato pZ130 cDNA-encoded protein.

KW Ovary; ovule; fruit; tomato; cotton; melanin; vector;

KM transgenic plant.

OS Lycopersicon esculentum cv. UC82B.

XX US5530185-A.

PD 25-JUN-1996.

```

XX PF 29-DEC-1992; 92US-0998158.
XX PR 29-DEC-1992; 92US-0998158.
XX PR 19-JUL-1989; 89US-0382518.
XX PR 17-JUL-1990; 90US-0554195.
XX PA (CALJ ) CALGENE INC.
XX PI Martineau BM, Reilley AA, Stalker DM;
XX DR N-PSDB; AAT31823.
XX PF 1996-308822/31.
XX DR N-PSDB; AAT31823.
XX PT DNA construct for expressing melanin synthesis gene in plant ovule
XX PT cells - contains promoter from the tomato p2130 gene, also binary
XX PT vector and transgenic plants, esp. cotton, contg. construct
XX PS Example 3; Fig 1A-B; 25pp; English.
XX CC The tomato p2130 clone contains a 564 bp insert of cDNA (AAT31823)
XX CC detected only in a tomato cDNA library prep. from pre-anthesis
XX CC RNA. It codes for a protein (AA97559) of unknown function. The
XX CC clone was isolated from a library prep. from cDNA of pre-anthesis
XX CC stage ovaries. The p2130 transcriptional initiation region is
XX CC considered to be ovary-specific and can be utilised in DNA
XX CC constructs for the expression of heterologous genes, partic. in
XX CC early fruit development.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX CC
XX SQ Sequence 118 AA;
XX
XX Query Match 60.3%; Score 333; DB 17; Length 118;
XX Best Local Similarity 62.5%; Pred. No. 3.3e-27;
XX Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;
OY 1 MARSICFMFAFALAMLFVAYEVQAR-CKTESNTPPGICITKPPCRKACISEKTTDHC 59
DB 14 MARSIFPMFALVAMMLFVTEVEAQQICKAPSQTFPGCLCMDSQCRKYCKEFTGHC 73
OY 60 SKILRCLCTKPCVDEKMTKTGAELAEAKTLAALBERIM 103
DB 74 SKLQRKCLCTKPCVFD-KISSEVKATLGSEAKTLSEVLEBERIM 116
XX
XX RESULT 5
XX ID AAM08364
XX AA AAM08364 standard; Protein; 118 AA.
XX AC AAM08364;
XX DT 14-MAR-1997 (first entry)
XX DE Tomato ovary-specific p2130-encoded protein.
XX KM Ovary; fruit; colour; pigmentation; cotton; tomato; probe;
XX KM promoter; p2130; p27; transgenic plant.
XX OS Lycopersicon esculentum cv. UC82B.
XX PN MO9640951-A2.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US09911.
XX PR 07-JUN-1995; 95US-0487087.
XX PA (CALJ ) CALGENE INC.
XX PI McBride K, Stalker DM;
XX DR WPI; 1997-052341/05.

```

```

XX DR N-PSDB; AAT48813.
XX PT DNA construct capable of directing ovary-tissue transcription in
XX PT plants - useful for modifying colour phenotype, in e.g. cotton
XX PS Example 1; Fig 1; 75pp; English.
XX CC Novel polypeptides (AAM08364 and AAM08385) are respectively encoded
XX CC by tomato p2130 (AAT48813) and p270 (AAT48814) ovary-specific genes.
XX CC Their native function is unknown. The promoter regions of the
XX CC genes can be used to direct ovary-specific transcription of a
XX CC gene of interest in transgenic plants.
XX SQ Sequence 118 AA;
XX
XX Query Match 60.3%; Score 333; DB 18; Length 118;
XX Best Local Similarity 62.5%; Pred. No. 3.3e-27;
XX Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;
OY 1 MARSICFMFAFALAMLFVAYEVQAR-CKTESNTPPGICITKPPCRKACISEKTTDHC 59
DB 14 MARSIFPMFALVAMMLFVTEVEAQQICKAPSQTFPGCLCMDSQCRKYCKEFTGHC 73
OY 60 SKILRCLCTKPCVDEKMTKTGAELAEAKTLAALBERIM 103
DB 74 SKLQRKCLCTKPCVFD-KISSEVKATLGSEAKTLSEVLEBERIM 116
XX
XX RESULT 6
XX ID AAR10310
XX AA AAR10310 standard; Protein; 179 AA.
XX AC AAR10310;
XX DT 25-MAR-2003 (updated)
XX DT 04-APR-1991 (first entry)
XX DE Ovary tissue transcriptional factor DNA clone p2130 product.
XX KM Ovary tissue transcriptional factor; DNA construct; probe;
XX KM clone p2130.
XX OS Lycopersicon esculentum UC82B.
XX PN EP409629-A.
XX PD 23-JAN-1991.
XX PF 19-JUL-1990; 90EP-0307926.
XX PR 19-JUL-1989; 89US-0382518.
XX PA (CALJ ) CALGENE INC.
XX PI Martineau B, Houck CM;
XX DR WPI; 1991-024191/04.
XX DR N-PSDB; AAO10262.
XX PT New ovary tissue transcriptional factors - modify transcription
XX PT in tomato plant ovaries for use as mol. probes
XX PS Disclosure; Fig 1; 21pp; English.
XX CC The tomato-derived transcriptional initiation region which regulates
XX CC the expression of the sequence corresp. to the p2130 clone is
XX CC considered ovary-specific. Sequences hybridisable to the p2130
XX CC clone, e.g. probe p27, show abundant mRNA, esp. at the early stages
XX CC of anthesis. The message is expressed in ovary integument and ovary
XX CC outer pericarp tissue and is not expressed, or at least not readily
XX CC detectable, in other tissues or at any other stage of fruit
XX CC development. The native function of the amino acid sequence
XX CC encoded by the structural gene comprising p2130 is unknown.

```

CC See also AA010263-64.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 XX Sequence 179 AA;  
 SQ  
 Query Match 60.3%; Score 333; DB 12; Length 179;  
 Best Local Similarity 62.5%; Pred. No. 5e-27;  
 Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;  
 Oy 1 MARSICFMAFAIILAMLFVAYEVOARE-CTESNTFPGICITKPPCKACISEKFTDGHG 59  
 Db 14 MARSIFMAFLVLAAMLFTYVEVAQOICRPSQTFPGICFMDSSCKRYCCKEFTGHC 73  
 Oy 60 SKILRCLCTKPCVFDKMTKGAETIAEAKTILAAALBEEIM 103  
 Db 74 SKLQRKCLCTKPCVFD-KISSEVKATLGEBAKTLSEVLEBEEIM 116  
 RESULT 7  
 AAR30779  
 ID AAR30779 standard; Protein: 188 AA.  
 XX  
 XX AAR30779;  
 XX AC  
 DT 25-MAR-2003 (updated)  
 DT 19-MAY-1993 (first entry)  
 XX  
 XX p2130 polypeptide.  
 XX cDNA; clone p2130; anthesis; tomato; ovary; integumen; outer pericarp;  
 KM fruit; development; transcription; initiation; region; modulation;  
 KM ovary-specific; endogenous; fruit product; exogenous; phenotype.  
 XX  
 OS Lycopersicon esculentum.  
 XX  
 XX Key Location/Qualifiers  
 FH Key  
 FT Misc-difference 119  
 FT /note= "Nonsense codon"  
 FT Misc-difference 120  
 FT /note= "Nonsense codon"  
 FT Misc-difference 126  
 FT /note= "Nonsense codon"  
 FT Misc-difference 129  
 FT /note= "Nonsense codon"  
 FT Misc-difference 146  
 FT /note= "Nonsense codon"  
 FT Misc-difference 152  
 FT /note= "Nonsense codon"  
 FT Misc-difference 160  
 FT /note= "Nonsense codon"  
 FT Misc-difference 164  
 FT /note= "Nonsense codon"  
 FT Misc-difference 181  
 FT /note= "Nonsense codon"  
 FT  
 XX US5175095-A.  
 PN  
 XX 29-DEC-1992.  
 PD  
 XX 17-JUL-1990; 90US-0554195.  
 PF  
 XX 19-JUL-1989; 89US-0382518.  
 PR 17-JUL-1990; 90US-0554195.  
 XX  
 XX (CALJ ) CALGENE INC.  
 PA  
 XX Houck CM, Martineau BM;  
 PI  
 XX WPI; 1993-026940/03.  
 DR N-PSDB; AAQ34940.  
 XX  
 PT DNA constructs contg. tomato p2130 transcriptional initiation  
 PT region - useful for modulation of endogenous fruit prods. and for

PT prodn. of exogenous prods.  
 XX  
 XX Disclosure; Fig 1; 18pp; English.  
 PS  
 XX The sequence represents the polypeptide of cDNA clone p2130. This  
 CC sequence is expressed during the early stages of anthesis in tomatoes.  
 CC The message is expressed in ovary integumen and ovary outer pericarp  
 CC tissue. It is not readily detectable in other tissues or at other  
 CC stages of fruit development. The transcription initiation region  
 CC associated with this gene is therefore considered to be ovary-  
 CC specific. The actual function of the p2130 polypeptide is unknown.  
 CC The transcription initiation region can be used for modulation of  
 CC endogenous fruit products, for production of exogenous products and  
 CC for modification of the phenotype of fruit and fruit products.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 XX  
 SQ Sequence 188 AA;  
 Query Match 60.3%; Score 333; DB 14; Length 188;  
 Best Local Similarity 62.5%; Pred. No. 5.3e-27;  
 Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;  
 Oy 1 MARSICFMAFAIILAMLFVAYEVOARE-CTESNTFPGICITKPPCKACISEKFTDGHG 59  
 Db 14 MARSIFMAFLVLAAMLFTYVEVAQOICRPSQTFPGICFMDSSCKRYCCKEFTGHC 73  
 Oy 60 SKILRCLCTKPCVFDKMTKGAETIAEAKTILAAALBEEIM 103  
 Db 74 SKLQRKCLCTKPCVFD-KISSEVKATLGEBAKTLSEVLEBEEIM 116  
 RESULT 8  
 ABP56255  
 ID ABP56255 standard; Protein: 188 AA.  
 XX  
 XX ABP56255;  
 AC  
 DT 28-MAR-2003 (first entry)  
 DT  
 XX Plasmid p2130 protein sequence p27.  
 DE  
 XX Transcriptional factor; ovary tissue.  
 KM  
 XX Unspecified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Key  
 FT Misc-difference 119  
 FT /label= "unknown"  
 FT Misc-difference 120  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 126  
 FT /label= "unknown"  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 129  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 146  
 FT /label= "unknown"  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 152  
 FT /label= "unknown"  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 160  
 FT /label= "unknown"  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 164  
 FT /label= "unknown"  
 FT /note= "encoded by a stop codon"  
 FT Misc-difference 181  
 FT /label= "unknown"  
 FT /note= "encoded by a stop codon"  
 FT

XX CN189856-A.  
 XX 05-AUG-1998.  
 XX 07-JUN-1996; 96CN-0195170.  
 XX 07-JUN-1995; 95US-0480087.  
 XX (CALJ) CALGENE INC.  
 XX MPI: 2002-733392/80.  
 XX N-PSDB; ABZ21970.  
 XX Use of transcriptional factors of ovary tissue -  
 XX Example 3; Fig 1A-B; 56pp; Chinese.  
 XX The present invention describes the use of transcriptional factors of  
 XX ovary tissue. The present sequence represents a protein sequence  
 XX which is used in an example from the present invention.  
 XX Sequence 188 AA;

Query Match 60.3%; Score 333; DB 23; Length 188;  
 Best Local Similarity 62.5%; Pred. No. 5.3e-27;  
 Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;

QY 1 MARSICFMALFALAMLFVAYEVQARE-CKTESNTPGICITKPPCKACISEKFTDGHG 59  
 DB 14 MARSIFPMALFVAMLFVAYEVQARE-CKTESNTPGICITKPPCKACISEKFTDGHG 73  
 QY 60 SKILRRCLCTKPCVDEKMTKTGAETLAEEAKTAAALAEIEM 103  
 DB 74 SKLQRKCLCTKPCVDF-KISSSVKATIGAEAKTISEVLAEEIEM 116

RESULT 9  
 AAB85079  
 ID AAB85079 standard; Protein; 84 AA.  
 XX AAB85079;  
 XX 22-AUG-2001 (first entry)  
 XX Pepper defensin protein (PepeDef) sequence.  
 XX KW Pepper; defensin; PepeDef; thionin-like protein; Pepthl; transgenic;  
 XX phytopathogen; antimicrobial.  
 XX Capsicum annum.  
 XX EP1101771-A1.  
 XX 23-MAY-2001.  
 XX 15-NOV-1999; 99EP-0309059.  
 XX 15-NOV-1999; 99EP-0309059.  
 XX (KOKU-) KOREA KUMHO PETROCHEMICAL CO LTD.  
 XX Oh B, Ko MK, Shin B, Chung CH;  
 XX MPI; 2001-357927/38.  
 XX N-PSDB; AAF83987.  
 XX New pathogen induced genes (pepper defensin protein gene and pepper  
 XX thionin-like protein gene) from Capsicum annum, useful for producing  
 XX transgenic plants with enhanced resistance against phytopathogens, e.g.  
 XX fungi or nematode -  
 XX Claim 1; Page 11; 22pp; English.

XX The invention provides new isolated nucleic acid molecules encoding a  
 CC pepper defensin protein (PepeDef) and a pepper thionin-like protein  
 CC (Pepthl). The PepeDef and Pepthl genes are useful for producing transgenic  
 CC plants that exhibit enhanced resistance against phytopathogens, e.g.  
 CC fungi, bacteria, viruses, nematode, mycoplasma-like organisms, parasitic  
 CC higher plants, flagellate protozoa or insects. The present sequence  
 CC represents the pepper defensin protein (PepeDef) sequence.  
 XX Sequence 84 AA;

Query Match 35.8%; Score 197.5; DB 22; Length 84;  
 Best Local Similarity 45.5%; Pred. No. 3.3e-13;  
 Matches 45; Conservative 9; Mismatches 28; Indels 17; Gaps 2;

QY 1 MARSICFMALFALAMLFVAYEVQARECKTESNTPGICITKPPCKACISEKFTDGHG 59  
 DB 1 MARSIFPMALFVAMLFVAYEVQARECKTESNTPGICITKPPCKACISEKFTDGHG 60  
 QY 60 SKILRRCLCTKPCVDEKMTKTGAETLAEEAKTAAAL 98  
 DB 61 FTLSKCLCKKRC-----NAXTLATELL 83

RESULT 10  
 ABU08325  
 ID ABU08325 standard; Protein; 84 AA.  
 XX ABU08325;  
 XX 29-MAY-2003 (first entry)

XX Capsicum annum L. cv. Hanbyul thionine.  
 XX Thionine; plant disease resistance; plant.  
 XX Capsicum annum L. cv. Hanbyul.

XX Key Location/Qualifiers  
 FT Misc-difference 5 /note= "Encoded by AAT"  
 FT FT Misc-difference 74 /label= "Xaa"  
 FT FT /note= "Xaa is given as "Net" in the specification,  
 and is encoded by AAT"

XX KR2002024732-A.  
 XX 01-APR-2002.  
 XX 26-SEP-2000; 2000KR-0056518.  
 XX 26-SEP-2000; 2000KR-0056518.  
 XX (KOCU-) KOREA CHUNGANG EDUCATIONAL FOUND.  
 XX Hwang BG, Kim YJ, Lee SC;  
 XX MPI; 2002-747901/81.  
 XX N-PSDB; ABX12800.  
 XX Thionine gene of capsicum annum 1. cv. hanbyul and probing method of  
 XX resistance for plant diseases -  
 XX Example 2; Fig 2; 14pp; Korean.  
 XX The present invention relates to the thionine gene of Capsicum  
 CC annum L. cv. Hanbyul, and a probing method of resistance for  
 CC plant diseases, caused by Xanthomonas campestris subsp. vesicatoria,  
 CC Collectotrichum coccodes, Collectotrichum gloeosporioides and the like.  
 CC The presents sequence represents Capsicum annum L. cv. Hanbyul  
 CC thionine.

SQ Sequence 84 AA;  
Query Match 34.5%; Score 190.5; DB 23; Length 84;  
Best Local Similarity 44.4%; Pred. No. 1.8e-12;  
Matches 44; Conservative 9; Mismatches 29; Indels 17; Gaps 2;  
CY 1 MARSLCPMAFATLARMFLVAVYQARECKTESNTPPGICITKPPCKRACI-SEKFTDGC 59  
DB 1 MARSTYMAFLVLAATYFVAVYQGKICKCKELTKPKCSSDPLCQKLCMEKEKEDGHC 60  
CY 60 SKILRRCLTKPCYFDEKMTKTGAELIAEFKTLAALL 98  
DB 61 FTLSKCLCMKRC-----XAKTLATELL 83  
RESULT 11  
ID AAG07868 standard; Protein; 100 AA.  
AC AAG07868;  
XX 17-OCT-2000 (first entry)  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 5184.  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5184.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
PN EP103405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 03-MAR-1999; 99US-0123160.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 26-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144884.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147132.

28.9%; Score 159.5; DB 21; Length 100;

QY	3	RSLCTMAVAILARMLFVAIEVQARECKTESNTFPeiCITTKPPCKRACISEKFTDGHCKSI	62
Db	32	RDSFVLVMTLATVW-GEVREAFARCTERTSNLFNGPCLSSNCANVCNNEGFSDDGCRGF	90
QY	63	LRRLCTKPC	72
Db	91	RRRLCTKRPC	100
RESULT 12			
ID	AA011734	standard; Protein, 128 AA.	
XX	AA011734		
AC	AA011734		
XX			
DT	17-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 10566.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway; hydridisation assay; genetic mapping; gene expression control; promoter; termination sequence.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	23-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	04-MAY-1999;	99US-0132407.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134321.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134370.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135553.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137528.	
PR	07-JUN-1999;	99US-0137724.	



```
PR 08-JUN-1999; 99US-0138034.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145091.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147206.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.9%; Score 159.5; DB 21; Length 128;
Best Local Similarity 44.3%; Pred. No. 4.8e-09;
Matches 31; Conservative 10; Mismatches 28; Indels 1; Gaps 1;
```

QY 3 RSLCFMFAILALRMFLVAYEQARECKTESNTFPGICITKPPORACISEKFTDGHCKSI 62  
DB 60 ROSFFVLFMLATVM-GPVRVEARTCTSSNLFNGPCLSSNCAVCHNEGFSDDDCNGF 118  
QY 63 LRRCCTKRC 72  
DB 119 RRRCLCTRPC 128

RESULT 13  
AAG25022  
ID AAG25022 standard; Protein; 76 AA.  
XX AAG25022;  
AC  
XX 17-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 28917.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141267.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 06-AUG-1999; 99US-0147803.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.

```
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148314.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158363.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160988.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161935.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.5%; Score 157.5; DB 21; Length 76;
Best Local Similarity 42.9%; Pred. No 4.5e-09;
Matches 30; Conservative 11; Mismatches 28; Indels 1; Gaps 1;
```

```
OY 3 RSLCFMFAIILAPMLFVAYEQABCKTESNTFPGICITKPPCKKACISKFTDGCSKI 62
DB 8 KSALLILFMILATVA-GPVRVEARTCTETSNTLFGNPGCLSSSNCANVCNIGGFSDDCGRF 66
```

```
OY 63 LRRLCTKPC 72
DB 67 RRRLCTRPC 76

RESULT 14
AAG07870
ID AAG07870 standard; Protein; 55 AA.
XX
AC AAG07870;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5186.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130444.
PR 23-APR-1999; 99US-0130810.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134371.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
```

Query	Match	Score	DB	Length
13-AUG-1999	99US-0148684	27.0%	21	55
16-AUG-1999	99US-0149368	25.0%	21	55
17-AUG-1999	99US-0149175	25.0%	21	55
18-AUG-1999	99US-0149426	25.0%	21	55
20-AUG-1999	99US-0149722	25.0%	21	55
20-AUG-1999	99US-0149723	25.0%	21	55
23-AUG-1999	99US-0149929	25.0%	21	55
23-AUG-1999	99US-0149902	25.0%	21	55
23-AUG-1999	99US-0149930	25.0%	21	55
26-AUG-1999	99US-0150566	25.0%	21	55
27-AUG-1999	99US-0150884	25.0%	21	55
27-AUG-1999	99US-0151065	25.0%	21	55
27-AUG-1999	99US-0151066	25.0%	21	55
30-AUG-1999	99US-0151080	25.0%	21	55
31-AUG-1999	99US-0151303	25.0%	21	55
01-SEP-1999	99US-0151438	25.0%	21	55
01-SEP-1999	99US-0151930	25.0%	21	55
07-SEP-1999	99US-0152363	25.0%	21	55
10-SEP-1999	99US-0153070	25.0%	21	55
13-SEP-1999	99US-0153758	25.0%	21	55
15-SEP-1999	99US-0154018	25.0%	21	55
16-SEP-1999	99US-0154039	25.0%	21	55
20-SEP-1999	99US-0154779	25.0%	21	55
22-SEP-1999	99US-0155119	25.0%	21	55
23-SEP-1999	99US-0155486	25.0%	21	55
24-SEP-1999	99US-0155659	25.0%	21	55
28-SEP-1999	99US-0156458	25.0%	21	55
29-SEP-1999	99US-0156596	25.0%	21	55
04-OCT-1999	99US-0157117	25.0%	21	55
05-OCT-1999	99US-0157753	25.0%	21	55
06-OCT-1999	99US-0157665	25.0%	21	55
07-OCT-1999	99US-0158029	25.0%	21	55
08-OCT-1999	99US-0158232	25.0%	21	55
12-OCT-1999	99US-0158359	25.0%	21	55
13-OCT-1999	99US-0159293	25.0%	21	55
13-OCT-1999	99US-0159294	25.0%	21	55
14-OCT-1999	99US-0159295	25.0%	21	55
14-OCT-1999	99US-0159330	25.0%	21	55
14-OCT-1999	99US-0159637	25.0%	21	55
14-OCT-1999	99US-0159638	25.0%	21	55
18-OCT-1999	99US-0159584	25.0%	21	55
21-OCT-1999	99US-0160741	25.0%	21	55
21-OCT-1999	99US-0160767	25.0%	21	55
21-OCT-1999	99US-0160768	25.0%	21	55
21-OCT-1999	99US-0160770	25.0%	21	55
21-OCT-1999	99US-0160814	25.0%	21	55
21-OCT-1999	99US-0160815	25.0%	21	55
22-OCT-1999	99US-0160980	25.0%	21	55
22-OCT-1999	99US-0160981	25.0%	21	55
22-OCT-1999	99US-0160989	25.0%	21	55
25-OCT-1999	99US-0161404	25.0%	21	55
25-OCT-1999	99US-0161405	25.0%	21	55
25-OCT-1999	99US-0161406	25.0%	21	55
26-OCT-1999	99US-0161359	25.0%	21	55
26-OCT-1999	99US-0161360	25.0%	21	55
26-OCT-1999	99US-0161361	25.0%	21	55
28-OCT-1999	99US-0161920	25.0%	21	55
28-OCT-1999	99US-0161992	25.0%	21	55
28-OCT-1999	99US-0161993	25.0%	21	55
29-OCT-1999	99US-0162142	25.0%	21	55

AA07869  
ID AA07869 standard; Protein: 61 AA.  
XX  
AC AA07869;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5185.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydrolisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126268.  
PR 29-MAR-1999; 99US-0126788.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139493.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144336.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0145986.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151103.  
 PR 01-SEP-1999; 99US-0151438.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0158293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 27.0%; Score 149; DB 21; Length 61;  
 Best Local Similarity 50.0%; Pred. No. 2.8e-08;  
 Matches 25; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

OY 23 VOARECKTESNTFFGICITKPPCKACISEKFTDGHCKIIRCLCTKPC 72  
 DB 12 VEARICETSSNLFNGPCLISSNVCANVCHNEGFSDDCGFRRCICTKPC 61

Search completed: January 28, 2004, 08:50:07  
 Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

# OM protein - protein search, using sw model

Run on: January 28, 2004, 08:46:51 ; Search time 20 Seconds  
(without alignments)  
504.886 Million cell updates/sec

Title: US-10-072-809B-18

Sequence: 1 MARSICFMAFALIAAMLFVA.....LAEAKTIAALLEEIMDN 105

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	529	95.8	105	2	S23574
2	333	60.3	105	2	S57809
3	150.5	27.3	79	2	T06381
4	149	27.0	55	2	D96636
5	147.5	26.7	81	2	T02667
6	144	26.1	78	2	S52634
7	143.5	26.0	73	2	D84433
8	138.5	25.1	74	2	S05994
9	137.5	24.9	77	2	S30578
10	137.5	24.9	77	2	C84433
11	127.5	23.1	77	2	B84433
12	125.5	22.7	77	2	T14395
13	119	21.6	82	2	UC7897
14	117.5	21.3	75	2	S51637
15	116	21.0	83	2	T14866
16	112	20.3	80	2	T02622
17	111	20.1	74	2	S24965
18	110	19.9	80	2	T10176
19	108	19.6	80	2	T10183
20	107	19.4	47	2	A58319
21	106.5	19.3	79	2	T02623
22	106	19.2	80	2	F96591
23	105	19.0	80	2	T02621
24	103	18.7	47	2	S69146
25	103	18.7	47	2	S69144
26	103	18.7	80	2	T10823
27	102.5	18.6	80	2	F96787
28	102.5	18.6	79	2	T07917
29	101	18.3	47	2	A58445

30	101	18.3	47	2	B58445	fabatin-2 - fava b
31	99	17.9	47	2	S13849	hordochionin gamma
32	99	17.9	48	2	S13963	alpha-amylase inhi
33	94.5	17.1	74	2	T06766	disease resistance
34	94.5	17.1	82	2	S65779	gamma-thionin prec
35	94	17.0	47	2	S11530	gamma-thionin gamma
36	92.5	16.8	75	2	S11536	gamma-purothionin
37	88	15.9	47	2	S11529	disease resistance
38	86.5	15.7	72	2	T06599	omega-hordochionin
39	84.5	15.3	48	2	S68952	plili protein (clon
40	83	15.0	83	2	T03673	dirosomycin precurs
41	77	13.9	70	2	A55824	gamma-zeachionin 2
42	76	13.8	47	2	B58319	salivary glue prot
43	73.5	13.3	297	2	S21085	salivary glue prot
44	72.5	13.1	264	2	S29893	gamma-thionin Si-a
45	72	13.0	47	2	S69145	

## ALIGNMENTS

RESULT 1  
S23574  
thionin precursor, flower-specific - common tobacco  
C.Species: Nicotiana tabacum (common tobacco)  
C.Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jun-2000  
C.Accession: S23574, S21000  
R.Gu, Q.; Kawata, E.E.; Morse, M.J.; Wu, H.M.; Cheung, A.Y.  
Mol. Gen. Genet. 234, 89-96, 1992  
A>Title: A flower-specific cDNA encoding a novel thionin in tobacco.  
A.Reference number: S23574; MUID:92357021; PMID:1495489  
A.Accession: S23574  
A.Molecule type: mRNA  
A.Residues: 1-105 <GU>  
A.Cross-references: EMBL:Z11748; NID:921212; PIDN:CAA77806.1; PID:S21213  
A.Experimental source: strain W38; tissue type flower buds  
A.Superfamily: gamma-thionin  
F.1-25/Domain: signal sequence #status predicted <SIG>  
F.26-105/Product: thionin #status predicted <MAP>  
F.28-72.39-59.45-66.49-68/Disulfide bonds: #status predicted

Query Match 95.8%; Score 529; DB 2; Length 105;  
Best Local Similarity 95.2%; Pred. NO. 7.6e-48;  
Matches 100; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	MARSICFMAFALIAAMLFVA	YVOARCKTESNTPFGICITKPCRCACISEKFTDGHCS	60
Db	1	MARSICFMAFALIAAMLFVA	YVOARCKTESNTPFGICITKPCRCACISEKFTDGHCS	60
Qy	61	KIARRCLCTKPCVFDEKMTGAE	ILAEAKTIAALLEEIMDN	105
Db	61	KIARRCLCTKPCVFDEKMTGAE	ILAEAKTIAALLEEIMDN	105

## RESULT 2

S57809  
gamma-thionin-like protein precursor - tomato  
N.Alternate names: A72 protein; defensin  
C.Species: Lycopersicon esculentum (tomato)  
C.Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jun-2000  
C.Accession: S57809; S72491  
R.Mulligan, S.B.; Gaeser, C.S.  
Plant Mol. Biol. 28, 691-711, 1995  
A>Title: Nature and regulation of p1scll-expressed genes in tomato.  
A.Reference number: S57808; MUID:95375233; PMID:7647301  
A.Accession: S57809  
A>Status: nucleic acid sequence not shown  
A.Molecule type: mRNA  
A.Residues: 1-105 <MTL>  
A.Cross-references: EMBL:U20591; NID:9224623; PIDN:AAA80496.1; PID:9224624  
A.Experimental source: flower; cv. VF36  
R.Brandstaedter, J.; Rossbach, C.; Theres, K.  
Mol. Gen. Genet. 252, 146-154, 1996

A>Title: Expression of genes for a defensin and a proteinase inhibitor in specific areas  
 A/Reference number: S72491; MUID:96397493; PMID:8804387  
 A/Accession: S72491  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-105 <BR>  
 A/Experimental source: shoot, cv. Moneymaker  
 C/Genetics:  
 A:Gene: AT2  
 C/Function:  
 A/Description: may respond to stimuli promoting the floral transition  
 C/Superfamily: gamma-thionin  
 F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:28-74/Product: gamma-thionin-like protein #status predicted <MAT>  
 F:75-105/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
 F:29-73,40-60,46-67,50-69/Disulfide bonds: #status predicted

Query Match 60.3%; Score 333; DB 2; Length 105;  
 Best Local Similarity 62.5%; Pred. No. 1.7e-27;  
 Matches 65; Conservative 12; Mismatches 25; Indels 2; Gaps 2;

QY 1 MARSLCFMAFALIAMLFVAYEVQAR-CKTESNTPPGICITKPPCRKACISEKPTDHC 59  
 DB 1 MARSLCFMAFALIAMLFVAYEVQAR-CKTESNTPPGICITKPPCRKACISEKPTDHC 60  
 QY 60 SKILRRCLCTPCVDFDECKTKTGAEILAEKTLAALLLEBEIM 103  
 DB 61 SKLQRKCLCTPCVDFD-KISEVYKATIGAEKTLSEVLEBEIM 103

RESULT 3  
 T06381  
 A:Proteinase inhibitor - soybean  
 C/Species: Glycine max (soybean)  
 C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jun-2000  
 C/Accession: T06381  
 R:Cushman, J.C.; Maitra, N.  
 A/Description: Isolation and expression of a drought-induced protease inhibitor from soy  
 A/Reference number: 215640  
 A/Accession: T06381  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-79 <CDS>  
 A/Cross-references: EMBL:U12150; NID:9533691; PIDN:AAC97524.1; PID:G533692  
 A/Experimental source: strain Essex; leaf  
 C/Superfamily: gamma-thionin

Query Match 27.3%; Score 150.5; DB 2; Length 79;  
 Best Local Similarity 36.7%; Pred. No. 1.2e-08;  
 Matches 29; Conservative 16; Mismatches 27; Indels 7; Gaps 1;

QY 1 MARSICFMAFALIAMLFVAYEV-----QAREKTESNTPPGICITKPPCRKACISEK 53  
 DB 1 MARSVFLVSTICVLILVATMGPTVVAEARTCESQSHRFGCLSDPTGCGVCRTER 60  
 QY 54 FTDGHSKILRRCLCTKPC 72  
 DB 61 FTGHCGRGFRRCFCCKHC 79

RESULT 4  
 D96636  
 A:Unknown protein, 87272-87105 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002  
 C/Accession: D96636  
 R/Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federisiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A66141; MUID:21016719; PMID:11130712  
 A/Accession: D96636

A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-55 <STO>  
 A/Cross-references: GB:AE005173; NID:96751695; PIDN:AAE27678.1; GSPDB:GN00141  
 C/Genetics:  
 A:Gene: TP1.20  
 A/Map position: 1  
 C/Superfamily: gamma-thionin

Query Match 27.0%; Score 149; DB 2; Length 55;  
 Best Local Similarity 50.0%; Pred. No. 1.2e-08;  
 Matches 25; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

QY 23 VQAREKTESNTPPGICITKPPCRKACISEKPTDHCSTILRRCLCTKPC 72  
 DB 6 VEARTEISSNLPNGPCSSNCAVCHNCHSDQCRGFRRCCLCTKPC 55

RESULT 5  
 T02667  
 A:Proteinase inhibitor - rice  
 C/Species: Oryza sativa (rice)  
 C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jun-2000  
 C/Accession: T02667  
 R:Yun, C.H.; Lee, J.H.; Park, J.H.; Lee, G.R.  
 A/Description: Molecular characterization of rice proteinase inhibitor gene.  
 A/Reference number: 214694  
 A/Accession: T02667  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-81 <YUN>  
 A/Cross-references: EMBL:AF044059; NID:G2829211; PIDN:AAC00503.1; PID:G2829212  
 A/Experimental source: strain IR36  
 C/Genetics:  
 A:Gene: RGP19  
 A/Insertions: 21/1  
 C/Superfamily: gamma-thionin

Query Match 26.7%; Score 147.5; DB 2; Length 81;  
 Best Local Similarity 37.0%; Pred. No. 2.4e-08;  
 Matches 27; Conservative 13; Mismatches 26; Indels 7; Gaps 1;

QY 7 FMAFALIAMLFVAYEV-----QAREKTESNTPPGICITKPPCRKACISEKPTDHC 59  
 DB 8 FSAAMLIMVLLIATGEMGPPVVAEARTCESQSHRFGCPARKANCAVCTEGFPDGTIC 67  
 QY 60 SKILRRCLCTKPC 72  
 DB 68 HGVRRCMCTKPC 80

RESULT 6  
 S52634  
 A:gamma-thionin - Petunia inflata  
 C/Species: Petunia inflata  
 C/Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jun-2000  
 C/Accession: S52634  
 R/Karunandaa, B.; Singh, A.; Rao, T.  
 Plant Mol. Biol. 26, 459-464, 1994  
 A/Title: Characterization of a predominantly pistil-expressed gene encoding a gamma-th  
 A/Reference number: S52634; MUID:95036017; PMID:7948992  
 A/Accession: S52634  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-78 <KAR>  
 A/Cross-references: GB:U27173; NID:9499654; PIDN:AAA4740.1; PID:9499655



C/Genetics: 22/1  
A/Introns: 22/1  
C/Superfamily: gamma-chionin

Query Match  
Best Local Similarity 35.9%; Pred. No. 5.4e-08;  
Matches 28; Conservative 14; Mismatches 30; Indels 6; Gaps 1;

Qy 1 MARSICFMAFALMLFVAYEV-----QARECKTESNTFGICITKPPCKACISEKFTDGS 54  
Db 1 MGRSIRLPATFPLMLFLSTEMGWTSAEARTCSGSHRHGTCVRSNCASVCQRTGTF 60  
Qy 55 TDGHSKILRCLCTKPC 72  
Db 61 IGGNCRFRRCCTTRNC 78

RESULT 7  
D84433  
protease inhibitor II [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
C/Accession: D84433  
R/Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vancken, S.E.; Umayam, L.; Tallon, L.;  
Euse, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: D84433  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-73 <STO>  
A/Cross-references: GB:A8002093; NID:94038038; PIDN:AAC97220.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: At2g02140  
A/Map position: 2  
C/Superfamily: gamma-chionin

Query Match  
Best Local Similarity 40.0%; Pred. No. 5.8e-08;  
Matches 30; Conservative 13; Mismatches 27; Indels 5; Gaps 2;

Qy 1 MARSICFMAFALMLFVAYEV-----EVOARECKTESNTFGICITKPPCKACISEKFTDGS 57  
Db 1 MKLSRLISALIMVLMFLATGMPVARTCESPSNFGVCLNSGCACACSESGSG 60  
Qy 58 HCSKILRCLCTKPC 72  
Db 61 RCSSL--RCYCSKAC 73

RESULT 8  
S05594  
pseudochionin St1 precursor - potato (strain cv. Bintje)  
C/Species: Solanum tuberosum (potato)  
A/Variety: strain cv. Bintje  
C/Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Jun-2000  
C/Accession: S05594; S45659  
R/Stekema, W.J.; Heidekamp, F.; Dirke, W.G.; van Beckum, J.; de Haar, P.; ten Bosch, C.  
Plant Mol. Biol. 11, 255-269, 1988  
A/Title: Molecular cloning and analysis of four potato tuber mRNAs.  
A/Reference number: S05592  
A/Accession: S05594  
A/Molecule type: mRNA  
A/Residues: 1-74 <STI>  
A/Cross-references: EMBL:X13190; NID:921393; PIDN:CAA1577.1; PTD:921394  
A/Experimental source: strain cv. Bintje  
A/Note: it is unknown whether 1-74 is the initiator or whether translation is initiated  
A/Note: the authors designated 1-74 as the protein as proteinase inhibitor (Bowman Birk) homoid  
Bur. J. Biochem. 223, 135-139, 1994  
A/Title: Pseudochionin-St1, a potato peptide active against potato pathogens.

A/Reference number: S45659; MUID:94307252; PMID:8033886  
A/Accession: S45659  
A/Molecule type: protein  
A/Residues: 28, N', 30-47 <MOR>  
A/Experimental source: strain cv. Desiree  
C/Superfamily: gamma-chionin  
P/1-27/Domain: signal sequence #status predicted <SIG>  
F/28-74/Product: pseudochionin St1 #status experimental <MAT>

Query Match  
Best Local Similarity 40.3%; Pred. No. 1.9e-07;  
Matches 29; Conservative 11; Mismatches 25; Indels 7; Gaps 2;

Qy 7 FNAFALMLFVAYEV-----QARECKTESNTFGICITKPPCKACISEKFTDGS 60  
Db 4 FATFFFLA-MLVATMGMRILAEARHCSLSHRKGPCTRRSNCASVETRFSGGNCH 62  
Qy 61 KILRCLCTKPC 72  
Db 63 GFARRCCTKPC 74

RESULT 9  
S30578  
protease inhibitor II [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Mar-2001  
C/Accession: S30578; A84433  
R/Yu, D.Y.; Quigley, F.; Maché, R.  
submitted to the EMBL Data Library, November 1992  
A/Description: Isolation and expression of a cDNA encoding protease inhibitor precursor  
A/Reference number: S30578  
A/Accession: S30578  
A/Residues: 1-77 <YUD>  
A/Molecule type: mRNA  
A/Cross-references: EMBL:X69139; NID:916426; PIDN:CAA4892.1; PID:916427  
R/Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vancken, S.E.; Umayam, L.; Tallon, L.;  
Euse, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: A84433  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-77 <STO>  
A/Cross-references: GB:A8002093; NID:94038041; PIDN:AAC97223.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: At2g02100  
A/Map position: 2  
C/Superfamily: gamma-chionin

Query Match  
Best Local Similarity 33.8%; Pred. No. 2.5e-07;  
Matches 26; Conservative 16; Mismatches 30; Indels 5; Gaps 1;

Qy 1 MARSICFMAFALMLFVAYEV-----EVOARECKTESNTFGICITKPPCKACISEKFTDGS 55  
Db 1 MKLSRLISALIMVLMFLATGMPVARTCESPSNFGVCLNSGCACACSESGSG 60  
Qy 56 DGHCKILRCLCTKPC 72  
Db 61 GGNCRGFRRCCTTRNC 77

RESULT 10  
C84433  
protease inhibitor II [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
C/Accession: C84433  
R/Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vancken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402:761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: C84433  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-77 <STO>  
 A:Cross-references: GB:AE002093; NID:G4038039; PIDN:AAC97222.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g02130  
 A:Map position: 2  
 C:Superfamily: gamma-thionin

Query Match 24.9%; Score 137.5; DB 2; Length 77;  
 Best Local Similarity 36.4%; Pred. No. 2.5e-07;  
 Matches 28; Conservative 15; Mismatches 29; Indels 5; Gaps 1;

QY 1 MARSLCFMAFATLARMLEFVA-----YEVQARECTESNTFPGICITKPPCKRACISEKFT 55  
 DB 1 MKLSVRFISALLFMVFATGMGPVTVEARTCESKSHRKPQCVSTHNCANVCNHEGFG 60

QY 56 DGHCKILRCLCTKPC 72  
 DB 61 GGRCKGRFRRCYCTRH 77

RESULT 11  
 B84433  
 proteinase inhibitor II [imported] - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
 C:Accession: B84433  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402:761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: B84433  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-77 <STO>  
 A:Cross-references: GB:AE002093; NID:G4038040; PIDN:AAC97222.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g02120  
 A:Map position: 2  
 C:Superfamily: gamma-thionin

Query Match 23.1%; Score 127.5; DB 2; Length 77;  
 Best Local Similarity 33.8%; Pred. No. 2.8e-06;  
 Matches 26; Conservative 12; Mismatches 34; Indels 5; Gaps 1;

QY 1 MARSLCFMAFATLARMLEFVA-----YEVQARECTESNTFPGICITKPPCKRACISEKFT 55  
 DB 1 MKFSMRLISAVLFLVMFVATGMGPVTVEARTCASQSRKGCVSPTNCENVCNHEGFG 60

QY 56 DGHCKILRCLCTKPC 72  
 DB 61 GGRCKGRFRRCYCTRH 77

RESULT 12  
 T14395  
 proteinase inhibitor II - turnip  
 C:Species: *Brassica rapa* (turnip)  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000  
 C:Accession: T14395  
 R:Kim, J.B.; Kim, H.U.; Yun, C.H.; Park, B.S.; Cho, W.S.; Ryu, J.C.; Chung, T.Y.  
 submitted to the EMBL Data Library, March 1996  
 A:Description: Nucleotide sequences of protease inhibitor II in chinese cabbage.  
 A:Reference number: Z18029  
 A:Accession: T14395

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-77 <KIM>  
 A:Cross-references: EMBL:L31937; NID:G1209257; PID:G1209258  
 A:Experimental source: unopened flower bud  
 C:Superfamily: gamma-thionin

Query Match 22.7%; Score 125.5; DB 2; Length 77;  
 Best Local Similarity 33.8%; Pred. No. 4.5e-06;  
 Matches 26; Conservative 14; Mismatches 32; Indels 5; Gaps 1;

QY 1 MARSLCFMAFATLARMLEFVA-----YEVQARECTESNTFPGICITKPPCKRACISEKFT 55  
 DB 1 MKLSMRLISAVLFLVMFVATGMGPVTVEARTCESKSHRKPQCVSTHNCANVCNHEGFG 60

QY 56 DGHCKILRCLCTKPC 72  
 DB 61 GGRCKGRFRRCYCTRH 77

RESULT 13  
 JG7897  
 defensin 1 precursor - wheat  
 C:Species: *Triticum aestivum* (common wheat)  
 C>Date: 03-Feb-2003 #sequence\_revision 03-Feb-2003 #text\_change 31-Mar-2003  
 C:Accession: JG7897  
 R:Koike, M.; Okamoto, T.; Tada, S.; Imai, R.  
 Biochem. Biophys. Res. Commun. 298, 46-53, 2002  
 A:Title: A novel plant defensin-like gene of winter wheat is specifically induced during  
 A:Reference number: JG7897; MUID:22266815; PMID:12379218  
 A:Accession: JG7897  
 A:Molecule type: mRNA  
 A:Residues: 1-82 <KO1>  
 A:Cross-references: DDBJ:AB089942  
 A:Experimental source: crown tissue  
 C:Comment: This protein is an antipathogenic protein and belongs to a subfamily of thic  
 d tolerance against pathogens during cold acclimation.  
 C:Genetics:  
 A:Gene: Tadi

Query Match 21.6%; Score 119; DB 2; Length 82;  
 Best Local Similarity 34.8%; Pred. No. 2.2e-05;  
 Matches 24; Conservative 13; Mismatches 24; Indels 8; Gaps 2;

QY 12 ILARMLFVAYEV-----QARECTESNTFPGICITKPPCKRACISEKFTDGHCKIL-- 63  
 DB 14 LLLMLLVATMGMTKTAARTCLSQSHKFGTCLSNCAVCRNTNFPDGBGNTLIVE 73

QY 64 RCLCTKPC 72  
 DB 74 RKCCKKRC 82

RESULT 14  
 S51637  
 probable proteinase inhibitor SE60 precursor - soybean  
 N:Alternate names: sulfur-rich protein SE60  
 C:Species: Glycine max (soybean)  
 C>Date: 01-Mar-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jun-2000  
 C:Accession: S51637; S51481  
 R:Lee, J.S.  
 submitted to the EMBL Data Library, July 1994  
 A:Reference number: S51637  
 A:Accession: S51637  
 A:Molecule type: DNA  
 A:Residues: 1-75 <LEE>  
 A:Cross-references: EMBL:Z18359; NID:G509768; PIDN:CAA79164.1; PID:G509769  
 A:Experimental source: strain Phangum  
 R:Choi, Y.; Ahn, U.H.; Choi, Y.D.; Lee, J.S.  
 Mol. Gen. Genet. 246, 266-269, 1995  
 A:Title: Tissue-specific and developmental regulation of a gene encoding a low molecular  
 A:Reference number: S51481; MUID:95166185; PMID:7862099  
 A:Accession: S51481

Query Match	21.3%	Score 117.5;	DB 2;	Length 75;
Best Local Similarity	33.3%	Pred. No. 3e-05;		
Matches 25; Conservative	14;	Mismatches 31;	Indels 5;	Gaps 2;

QY 1 MARSLCEFAAFILIRMF--VAEYVOARECKTESNFPGICITKPCFRCACISSEKPTGG 5  
 Db 3 MRKSGGF--FLLILLVYFASGVYVGTBGRVCESSHGFHGLCNRDHNCALVCGNEGSGG 6  
 QY 58 HCSKILRRCLCTKPC 72  
 Db 61 RCKGFRRRCPCCTRIC 75

## RESULT 15

Probable gamma-thionin precursor SP1 - Norway spruce  
C/Species: Picea abies (Norway spruce)  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Jun-2000  
C/Accession: T14866  
R/Sharma, P., Limeborg, A.  
Submitted to the EMBL Data Library, September 1995  
A/Description: Isolation and characterization of a cDNA encoding a gamma-thionin-like protein  
A/Reference number: Z18247  
A/Accession: T14866  
A/Status: preliminary; translated from GB/EMBL/DBD  
A/Molecule type: mRNA  
A/Residues: 183 <SHA>  
A/Cross-references: EMBL:X91487; NID:G1360107; PID:e198939  
A/Experimental source: root  
C/Genetics:  
A/Gene: SP1  
C/Superfamily: gamma-thionin  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-83/Product: probable gamma-thionin precursor #status predicted <MAT>

Query Match	21.0%;	Score 116;	DB 2;	Length 83;
Best Local Similarity	39.2%;	Pred. No. 4.6e-05;		
Matches 20; Conservative	8;	Mismatches 21;	Indels 2;	Gaps 1.

D5  
QY 24 QARECKTESNTFPGICITKPPCRACISBEFTDGHGKIL--RRLCTKPC 72  
          : | | | : | : | : | : | : | : | : | : |  
Db 32 EGRCTKTPSGKFGVCASSNNCKNVCCQTGEFPSGSCEFHVANKCYCSKPC 82

Search completed: January 28, 2004, 08:48:40  
Job time : 21 secs

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

11286.134 Million cell updates/sec

Sequence: 1 atggctcgcctccctgtgctt.....tgggttaaaaaaaaaa 541

[illegible]

Total number of hits satisfying chosen parameters: 5777422

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

```

1:  GenEmb1:*
2:  gb_ba:*
3:  gb_htg:*
4:  gb_in:*
5:  gb_om:*
6:  gb_ov:*
7:  gb_pat:*
8:  gb_ph:*
9:  gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_scs:*
13: gb_sy:*
14: gb_un:*
15: gb_vl:*
16: em_ba:*
17: em_fun:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

Result	Score	Query	Match	Length	DB	ID	Description
1	539.4	99.7	601	8	AF509666	AF509666	Nicotiana
2	504.2	93.2	589	8	SESTKXNA	SESTKXNA	Z11748 Solanaceae
3	482.8	89.2	558	8	AB005266	AB005266	Nicotiana
4	370.8	68.5	549	8	AB005250	AB005250	Nicotiana
5	354.8	65.6	448	8	AB005265	AB005265	Nicotiana
6	191.4	35.4	596	8	SLU02091	SLU02091	Solanum lyc
7	163	30.1	559	8	AF128239	AF128239	Capsicum
8	133.8	24.7	456	8	AF507975	AF507975	Petunia x
9	122.2	22.6	602	8	AF507976	AF507976	Petunia x
10	95.6	17.7	506	8	AF112443	AF112443	Capsicum
11	95.6	17.7	548	8	AF112869	AF112869	Capsicum
12	49.2	9.1	126366	9	HS104383	HS104383	Human DNA
13	49	9.1	172706	2	AC134847	AC134847	Human DNA
14	48.6	9.0	156060	2	AC0004153	AC0004153	Mus muscu
15	48.6	9.0	250707	3	AE014848	AE014848	Plasmodiu
16	48.4	8.9	1335	1	AF182282	AF182282	Buchnera
17	48	8.9	114694	9	AI355500	AI355500	Buchnera
18	48	8.9	349980	6	AX344550	AX344550	Human DNA
19	47.8	8.8	47296	3	AF160864	AF160864	Tetrahyme
20	47.4	8.8	3361	3	AF482383	AF482383	Dictyoste
21	47.4	8.8	19634	6	AX598816	AX598816	Sequence
22	47.4	8.8	19634	6	AX598962	AX598962	Sequence
23	46.2	8.5	1329	1	AF182262	AF182262	Buchnera
24	46.2	8.5	1330	1	AF182253	AF182253	Buchnera
25	46	8.5	57535	1	AC115682	AC115682	Dictyoste
26	45.8	8.5	1335	1	AF182280	AF182280	Buchnera
27	45.8	8.5	40780	3	CEM03C11	CEM03C11	Buchnera
28	45.4	8.4	5016	3	DDU31631	DDU31631	Dictyoste
29	45.4	8.4	4775	3	AF479338	AF479338	Dictyoste
30	45.4	8.4	7218	6	166494	166494	Sequence 14
31	45.4	8.4	237746	2	AC125778	AC125778	Sequence 14
32	45.4	8.4	271932	2	AC097673	AC097673	Rattus no
33	45.4	8.4	279698	2	PFM14P4	PFM14P4	Plasmodiu
34	45.4	8.4	348600	1	AB065521	AB065521	Wiggleswo
35	45.2	8.4	1335	1	AY182276	AY182276	Buchnera
36	45.2	8.4	1335	1	AY182281	AY182281	Buchnera
37	45.2	8.4	13363	3	MSQWCG1	MSQWCG1	Buchnera
38	45.2	8.4	34548	6	AX349035	AX349035	Anopheles g
39	45.2	8.4	167136	2	AC116305	AC116305	Dictyoste
40	45.2	8.4	167136	2	BX537323	BX537323	Danio rer
41	45.2	8.4	266544	3	AC116956	AC116956	Dictyoste
42	44.8	8.3	212927	2	BX248115	BX248115	Danio rer
43	44.6	8.2	1301	1	AY182279	AY182279	Buchnera
44	44.6	8.2	1330	1	AY182252	AY182252	Buchnera
45	44.6	8.2	1330	1	AY182254	AY182254	Buchnera

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AF509566	AF509566	Nicotiana glauca flower-specific defensin precursor (Nad1) mRNA, complete cds.	AF509566	AF509566.1	GI:25005096	Nicotiana glauca (Persian tobacco)	Nicotiana glauca	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.	Lay, P.T., Schirra, H.J., Scanlon, M.J., Anderson, M.A. and Craik, D.J.

TITLE	The three-dimensional solution structure of Nabl, a new floral defensin from <i>Nicotiana glauca</i> and its application to a homology model of the crop defense protein alfalfa
JOURNAL	J. Mol. Biol. 325 (1), 175-188 (2003)
MEDLINE	22361678
PUBMED	12472460
REFERENCE	2 (bases 1 to 601)
AUTHORS	Lay, F.T., Bruguiera, F. and Anderson, M.A.
TITLE	Isolation and Properties of Floral Defensins from Ornamental Tobacco and Petunia
JOURNAL	Plant Physiol. 131 (3), 1283-1293 (2003)
PUBMED	12644678
REFERENCE	3 (bases 1 to 601)
AUTHORS	Lay, F.T. and Anderson, M.A.
TITLE	Direct Submission
JOURNAL	Submitted (06-MAY-2002) Biochemistry, La Trobe University, Bundoora, VIC 3086, Australia
FEATURES	Location/Qualifiers

```

gene
1..601
/gene="Nad1"
CDS
61..378
/gene="Nad1"
/note="Nad1, gamma-thionin; possesses antifungal activity
targeted to the vacuole"
/codon_start=1
/product="flower-specific defensin precursor"
/protein_id="AA107999.1"
/db_xref="GI:25005037"
/translation="MARSLCFMFAIILAMLVFAVEYQARECTSEINTPGICIKP
CRACISEKFTDGHOSKILRLRLCTKPCVDENKTKGAILAEAKTLAALLDEE
MN"
mat_peptide
136..276
/gene="Nad1"
/product="flower-specific defensin"
BASE COUNT      192 a      95 c      115 g      199 t
ORIGIN

```

Query Match	99.7%	Score 539.4	DB 8	Length 601
Best Local Similarity	99.8%	Pred. No. 6.3e-102		
Matches 540	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 1	ATGGCTGCCTCTTGCTTCA	TGGCATTTGCTATCTTTGGCA	GGATGCTCTTTGTGGC	60
Db 61	ATGGCTGCCTCTTGCTTCA	TGGCATTTGCTATCTTTGGCA	GGATGCTCTTTGTGGC	120
QY 61	TATGAGTGCAGCTTGAGAA	TGCACAAACAGAAAGCAAC	CACTTCTGGAAATATG	120
Db 121	TATGAGTGCAGCTTGAGAA	TGCACAAACAGAAAGCAAC	CACTTCTGGAAATATG	180
QY 121	ACCAAAACACATGCAGAAA	AGCTTGATGATGAGAAATTA	CTGATGCTCATTTAGC	180
Db 181	ACCAAAACACATGCAGAAA	AGCTTGATGATGAGAAATTA	CTGATGCTCATTTAGC	240
QY 181	AAAATCCTGAGAAAGTGCT	TATGATCAAGCAATGTGTT	TGATGAGAAATGACT	240
Db 241	AAAATCCTGAGAAAGTGCT	TATGATCAAGCAATGTGTT	TGATGAGAAATGACT	300
QY 241	ACAGAGAGCTGAAATTTTG	CTGAGGAACAAAACCTTTG	GTGAGCTTGCTGAAGA	300
Db 301	ACAGAGAGCTGAAATTTTG	CTGAGGAACAAAACCTTTG	GTGAGCTTGCTGAAGA	360
QY 301	GAGATTAATGGATTAATTA	GATTAAGATTAAGAAATTA	AGATGACATACACAAAT	360
Db 361	GAGATTAATGGATTAATTA	GATTAAGATTAAGAAATTA	AGATGACATACACAAAT	420
QY 361	AAAGTTTCACTTCTTTTAA	AGGTGACTTAATGTGTGTT	TATATGGCTTTATATAC	420
Db 421	AAAGTTTCACTTCTTTTAA	AGGTGACTTAATGTGTGTT	TATATGGCTTTATATAC	480

Qy	421	CTTTATATACCTTTAAATTAAGTGGGACCTCAATCCTTTGGCAATCTTGACATPAGT	480
Db	481	CTTTATATACCTTTAAATTAAGTGGGACCTCAATCCTTTGGCAATCTTGACATPAGT	540
Qy	481	TTATTTGGTACTCTTTAATGAAAAAGACCTCTATGGCTTTGGTTAAAAA	540
Db	541	TTATTTGGTACTCTTTAATGAAAAAGACCTCTATGGCTTTGGTTAAAAA	600
Qy	541 A	541	
Db	601 A	601	

RESULT 2	589 bp	MRNA	linear	PLN 03-DEC-1993
SEQUENCE	Solanaece mRNA for flower-specific thionin.			
LOCUS	211748.1	541631		
DEFINITION	Zm1748.1 GI:21212			
ACCESSION	Zm1748.1			
VERSION	1			
KEYWORDS	thionin.			
SOURCE	Solanaece			
ORGANISM	Solanaceae			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Asteridae; lamids; Solanales.			
REFERENCE	1 (bases 1 to 589)			
AUTHORS	Gu,Q., Kawata,E.R., Morse,M.J., Wu,H.M. and Cheung,A.Y.			
TITLE	A flower-specific cDNA encoding a novel thionin in tobacco			
JOURNAL	Mol. Gen. Genet. 234 (1), 89-96 (1992)			
URL	92357021			
MEDLINE	1495489			
REFERENCE	2 (bases 1 to 589)			
AUTHORS	Gu,Q.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-FEB-1992) Qing GU, Department of Biology, Yale University, 165 Prospect St, New Haven, CT, 06511, USA			
FEATURES	Location/Qualifiers			
source	1..589			

	CDS	ORIGIN
	/db_xref="taxon:4070" /clone="49A" /tissue_type="flower of tobacco" /dev_stage="flowering" 49.._366 /codon_start=1 /product="flower-specific thionin" /protein_id="CAI77806.1" /db_xref="GI:21213" /db_xref="SWISS-PROT:P32026" /translation="MARSLCFMFAILLMMFLFVAVHVOARECKTESNTPFGICITKPKP CRKAISKEKFTDGCSTLRLRCCTKRFQVEKMIKTGAETLVEAKTLAAALDEEBI MDN"	184 a      97 c      114 g      194 t
Query Match	93.2%	Score 504.2; DB 8;
Best Local Similarity	95.7%;	Pred. No. 1.3e-94;
Matches 518;	Conservative 0;	Mismatches 23; Indels 0; Gaps 0.
Dy	1 ATGGCTGCTCCTTGTCCTTCATGGCATTTGCTATCTTGCGAAGATGCTCTTTGTGCC	60
Dd	49 ATGGCTGCTCCTTGTCCTTCATGGCATTTGCTATCTTGCGAATGATGCTGTTGTGCC	108
Dy	61 TATGAGGTCAAGTAGAGAATGACAAAACAAGAACAACATTTCTCTGGAATATGACTT	120
Dd	109 TATGAGGTCAAGTAGAGAATGACAAAACAAGAACAATCATTTCTCTGGAATATGACTT	168
Dy	121 ACCAAACCACATGAGAGAAAGCTTGTATATGAGAGAAATTATCTGATATGTATGTAC	180
Dd	169 ACCAAACCACATGAGAGAAAGCTTGTATATGAGAGAAATTATCTGATATGTATGTAC	228
Dy	181 AAAATCTCAGAGGTGCCTATGTACTTAAGCATGTGTGTTTGATGAGAAAGATGACTAAA	240

Db 229 AAACCTCTGACGCGTGCCTTATGCTAGCCATGCTGTGTATGATGAGAAATGATATAA 288

QY 241 ACAGAGCTGAAATTTTGGCTGAGAGCAAAAACTTTGGCTGACCTTTGCTTGAAGA 300

Db 289 ACAGAGCTGAAATTTTGGCTGAGAGCAAAAACTTTGGCTGACCTTTGCTTGAAGA 348

QY 301 GAGATATGATGATTAATTAAGATTAAGAAATTAAGATGACATTCACATTAAT 360

Db 349 GAGATATGATGATTAATTAAGATTAAGAAATTAAGATGACATTCACATTAAT 408

QY 361 AAAGTTTCTACCTTTCTTAAAGTGTAGCTAAATGTTGTATTAATTTGCTTTAGTAC 420

Db 409 AAAGTTTCTACCTTTCTTAAAGTGTAGCTAAATGTTGTATTAATTTGCTTTAGTAC 468

QY 421 CTTTATTAACCTTTAAATTAAGTGTGACCTTCAATCTTTGTGCAATCTTGCACTAAGT 480

Db 469 CTTTGTGTACCTTTAAATTAAGTGTGACCTTCAATCTTTGTGCAATCTTGCACTAAGT 528

QY 481 TTATTTGTGACTTTTAATGAATAATGACCTTGTATGCTTTGTTAAATTAATTAATTAAT 540

Db 529 TTATTTGTGACTTTTAATGAATAATGACCTTGTATGCTTTGTTAAATTAATTAATTAAT 588

QY 541 A 541

Db 589 A 589

RESULT 3  
AB005266 558 bp mRNA linear PLN 05-FEB-1999  
LOCUS Nicotiana excelsior mRNA for gamma-thionin, complete cds.  
DEFINITION  
AB005266  
VERSION  
AB005266.1 GI:2244704  
KEYWORDS  
gamma-thionin  
SOURCE  
Nicotiana excelsior  
Nicotiana excelsior  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Nicotiana.  
1 (sites)  
Yamada, S., Komori, T. and Imaseki, H.  
CDNA Cloning of gamma-thionin from Nicotiana excelsior (Accession  
No. AB005266) (PGR97-131)  
Plant Physiol. 115, 314 (1997)  
2 (bases 1 to 558)  
Yamada, S.  
Direct Submission  
Submitted (26-JUN-1997) Shigehiro Yamada, Plant Breeding and  
Genetics Research Laboratory, Japan Tobacco Inc., 700 Higashibara,  
Iwata, Shizuoka 438, Japan  
(E-mail: Shigehiro.Yamada@pbgrl.jti.co.jp, Tel:0538-32-7116,  
Fax:0538-32-8700)

FEATURES  
source  
1..558  
/organism="Nicotiana excelsior"  
/mol\_type="mRNA"  
/db\_xref="taxon:61185"  
1..558  
/gene="Nethio2"  
33..350  
/gene="Nethio2"  
/codon\_start=1  
/product="gamma-thionin"  
/protein\_id="BA2114.1"  
/db\_xref="GI:2244705"  
/translation="MARSVCFMAFAIILAVLVAVDVADCKTESNTPGICITKRP  
CRKACIKKFTDCHSKILRCLCTKPCVFEKMKITGAETLAEATTLAALLEEI  
MDN"

BASE COUNT 170 a 94 c 108 g 186 t

ORIGIN

Query Match 89.2%; Score 482.8; DB 8; Length 558;

Best local similarity 94.9%; Pred. No. 3.5e-90;  
Matches 499; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATGGCTGCTCCTTGGCTTCATGCAATCTTGCAAGATGCTCTTTGTTGCC 60

Db 33 ATGGCTGCTCCTTGGCTTCATGCAATCTTGCAAGATGCTCTTTGTTGCC 92

QY 61 TATGAGTGCAGAGTGAAGATGCAAAACAGAAAGCAACATTTCCGTGAATATGCAAT 120

Db 93 TATGATGGAAGCTAAAGATGCAAAACAGAAAGCAATATTCCTGGAAATATGAT 152

QY 121 ACCAACCACCAAGCAAAAAAGCTGTATGAGAAATTTACTGATGCTTGTAGC 180

Db 153 ACCAACCACCAAGCAAAAAAGCTGTATGAGAAATTTACTGATGCTTGTAGC 212

QY 181 AAATCTCAGAAAGTGCCTATGATCAATGACCATGTGTGTTATGAGAAATGACATAA 240

Db 213 AAATCTCAGAAAGTGTATGATCAATGACCATGTGTGTTATGAGAAATGACATAA 272

QY 241 ACAGAGCTGAAATTTTGGCTGAGAGCAAAAACTTTGGCTGACCTTTGCTTGAAGA 300

Db 273 ACAGAGCTGAAATTTTGGCTGAGAGCAAAAACTTTGGCTGACCTTTGCTTGAAGA 332

QY 301 GAGATATGATGATTAATTAAGATTAAGAAATTAAGATGACATTCACATTAAT 360

Db 333 GAGATATGATGATTAATTAAGATTAAGAAATTAAGATGACATTCACATTAAT 382

QY 361 AAAGTTTCTACCTTTCTTAAAGTGTAGCTAAATGTTGTATTAATTTGCTTTAGTAC 420

Db 393 AAAGTTTCTACCTTTCTTAAAGTGTAGCTAAATGTTGTATTAATTTGCTTTAGTAC 452

QY 421 CTTTATTAACCTTTAAATTAAGTGTGACCTTCAATCTTTGTGCAATCTTGCACTAAGT 480

Db 453 CTTTATTAACCTTTAAATTAAGTGTGACCTTCAATCTTTGTGCAATCTTGCACTAAGT 512

QY 481 TTATTTGTGACTTTTAATGAATAATGACCTTGTATGCTTTGTTT 526

Db 513 TTATTTGTGACTTTTAATGAATAATGACCTTGTATGCTTTGTTT 558

RESULT 4  
AB005250 549 bp mRNA linear PLN 19-JUN-1999  
LOCUS Nicotiana paniculata mRNA for gamma-thionin, complete cds.  
DEFINITION  
AB005250  
VERSION  
AB005250.1 GI:2251080  
KEYWORDS  
gamma-thionin  
SOURCE  
Nicotiana paniculata  
Nicotiana paniculata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; Lamiales; Solanales; Solanaceae; Nicotiana.  
1 (sites)  
Komori, T., Yamada, S. and Imaseki, H.  
A CDNA clone for gamma-thionin from Nicotiana paniculata (Accession  
No. AB005250) (PGR97-132)  
Plant Physiol. 115, 314 (1997)  
2 (bases 1 to 549)  
Komori, T.  
Direct Submission  
Submitted (26-JUN-1997) Toshiyuki Komori, Plant Breeding and  
Genetics Research Laboratory, Japan Tobacco Inc., 700 Higashibara,  
Iwata, Shizuoka 438, Japan  
(E-mail: Toshiyuki.Komori@pbgrl.jti.co.jp, Tel:0538-32-7116,  
Fax:0538-32-8700)

FEATURES  
source  
1..549  
/organism="Nicotiana paniculata"  
/mol\_type="mRNA"  
/db\_xref="taxon:62141"  
1..549  
/gene="NpThio1"  
48..368

gene

CDS







Db 336 CTGAGAGAGAGATTCCTGCGGCTGTAATTAAGTCTGATTAGATTAGTGT---CACAC 392  
 QY 352 ACACATATAAAGTTTACCTTTCTTAAGAAGTACGTAATGTTGGTTTAAATGGCT 411  
 Db 393 AAAATTAAATATATGTGTGCTTTCTTAAGAAGTACGTAATGTTGGCT 452  
 QY 412 TTATAGACCTTTTATTAACATTAAATAGTGGACCTCAATCTTTGSCAAATCTT 471  
 Db 453 TATAGTAGCATTGACACACA-TTAAATAGTGTGACACATAAATCTATATATGATCTTG 511  
 QY 472 GCAGTATTTATTTGTGTACTTTTATGAAGAAATGAC 508  
 Db 512 TATTAAGTTTATGTGTTTATATGAAAAGACCGCTC 548

RESULT 8  
 AF507975 456 bp mRNA linear PLN 19-MAR-2003  
 LOCUS Petunia x hybrida floral defensin-like protein 1 (D1) mRNA,  
 DEFINITION complete cds.  
 ACCESSION AF507975.1 GI:24817407  
 VERSION AF507975.1  
 KEYWORDS  
 SOURCE Petunia x hybrida  
 ORGANISM Petunia x hybrida  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; lamids; Solanales; Solanaceae; Petunia.

REFERENCE  
 AUTHORS Lay, F.T., Brugliera, F. and Anderson, M.A.  
 TITLE 1 (bases 1 to 456)  
 JOURNAL Isolation and Properties of Floral Defensins from Ornamental  
 REFERENCE Plant Physiol. 131 (3), 1283-1293 (2003)  
 PUBMED 12644678  
 AUTHORS Brugliera, F., Holton, T.A., Stevenson, T. and Mason, J.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-APR-2002) Florigene Ltd, 16 Gips St., Collingwood,  
 Melbourne, Victoria 3066, Australia

FEATURES  
 source location/Qualifiers  
 1..456  
 /organism="Petunia x hybrida"  
 /mol\_type="mRNA"  
 /cultiivar="Old Glory Blue"  
 /db\_xref="taxon:4102"  
 gene 1..456  
 /gene="D1"  
 /note="Phd1"  
 CDS 68..379  
 /gene="D1"  
 /codon\_start=1  
 /product="floral defensin-like protein 1"  
 /protein\_id="AA064750.1"  
 /db\_xref="GI:24817408"  
 /translation="MARSLCPFAVAIIAALMLFAAYDAEAATCKACBPWDSVCINKPK  
 CVACKCKAKSDHCHSKILRCLCTKECEVEKTEATOTETFTDVNTTALALIEADMM  
 V"  
 BASE COUNT 124 a 92 c 102 g 138 t  
 ORIGIN

Query Match 24.7%; Score 133.8; DB 8; Length 456;  
 Best Local Similarity 65.0%; Pred. No. 9e-18;  
 Matches 214; Conservative 0; Mismatches 112; Indels 3; Gaps 1;

QY 1 ATGGCTGCTCTTCTTGCTTCATGGCAATTTGCTATCTTGSCAAGATGCTCTTTGTTGCC 60  
 Db 68 ATGGCTGCTCCACATCTGTTCTTTCGCGGTGCTAATACATGATGATGCTCTTTGCTGCC 127  
 QY 61 TATGAGTGAAGCTAGAGATGCAAAACAGAAAGCAACATTTCTGGAATATGCAATT 120  
 Db 128 TATGAGCAGAAAGGCGCACTTGCAAGGCTGAAATGCCCAATGCGATTCGCTGTATC 187  
 QY 121 ACCAAACCAACCATSCAGAAAAGCTTGATACAGTGAAGAAATTTACTAGTGTCATGTAGC 180

Db 188 AATAAAAAACCATGTGTACGCTGTGTGCAAAAAGCAAAAGTCTCTGATGGGCACTGACGC 247  
 QY 181 AAAATCCTCAGAAAGTGGCTATATGATCAAGCAATGCTGTGTTGATGGAATGATCAATAA 240  
 Db 248 AAAATACCTCAGAAAGTGTGTGTGACCTAAGAGTGTATTCGAAAAAATGAGGC---A 304  
 QY 241 ACAGAGCTGAATTTTGGCTGAGAGCAAAAATTTGGCTGACGCTTTGCTTGAAGAA 300  
 Db 305 ACTCAAAACAGAAATTTCTACTAGAGATGTAAACATTTGGCTGAAGCTTTGCTTGAAGCA 364  
 QY 301 GAGATATGATGATTAATTAGATTAG 329  
 Db 365 GATATGATGTGTAGAGGTTACTTGTTAG 393

RESULT 9  
 AF507976 602 bp mRNA linear PLN 19-MAR-2003  
 LOCUS Petunia x hybrida floral defensin-like protein 2 (D2) mRNA,  
 DEFINITION complete cds.  
 ACCESSION AF507976  
 VERSION AF507976.1 GI:24817409  
 KEYWORDS  
 SOURCE Petunia x hybrida  
 ORGANISM Petunia x hybrida  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; lamids; Solanales; Solanaceae; Petunia.

REFERENCE  
 AUTHORS Lay, F.T., Brugliera, F. and Anderson, M.A.  
 TITLE 1 (bases 1 to 602)  
 JOURNAL Isolation and Properties of Floral Defensins from Ornamental  
 REFERENCE Plant Physiol. 131 (3), 1283-1293 (2003)  
 PUBMED 12644678  
 AUTHORS Brugliera, F., Herbert, S., Stevenson, T. and Mason, J.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-APR-2002) Florigene Ltd, 16 Gips St., Collingwood,  
 Melbourne, Victoria 3066, Australia

FEATURES  
 source location/Qualifiers  
 1..602  
 /organism="Petunia x hybrida"  
 /mol\_type="mRNA"  
 /cultiivar="Old Glory Blue"  
 /db\_xref="taxon:4102"  
 gene 1..602  
 /gene="D2"  
 /note="Phd2"  
 CDS 67..372  
 /gene="D2"  
 /codon\_start=1  
 /product="floral defensin-like protein 2"  
 /protein\_id="AA064751.1"  
 /db\_xref="GI:24817410"  
 /translation="MARSLCPFAVAIIAALMLFAAYETEAATCKACBPWDSVCINKPK  
 CVCKCKAPKPEFTDGHCHSKILRCLCTKECEVEKTEATOTETFTDVNTTALALIEADMM  
 V"  
 BASE COUNT 221 a 102 c 114 g 165 t  
 ORIGIN

Query Match 22.6%; Score 122.2; DB 8; Length 602;  
 Best Local Similarity 67.2%; Pred. No. 2.1e-15;  
 Matches 213; Conservative 0; Mismatches 83; Indels 21; Gaps 2;

QY 1 ATGGCTGCTCTTCTTGCTTCATGGCAATTTGCTATCTTGSCAAGATGCTCTTTGTTGCC 60  
 Db 67 ATGGCTGCTCCACATCTGTTCTTTCGCGGTGCTAATACATGATGATGCTCTTTGCTGCC 126  
 QY 61 TATGAGTGAAGCTAGAGATGCAAAACAGAAAGCAACATTTCTGGAATATGCAATT 120  
 Db 127 TATGAGCAGAAAGGCGCACTTGCAAGGCTGAAATGCCCAATGCGATTCGCTGTATC 186  
 QY 121 ACCAAACCAACCATSCAGAAAAGCTTGTA-----TCAGTGAAGAAATTTACTAGTGTCAT 174

Db 187 AATAAGCCCATGTGTAAATGTGTGCAAGACACCAAGAAATTCACCGAGGGGAC 246  
 QY 175 TGTACCAAAATGCTGAGAGGTCCTATGTACTAGCATGCTGTTGATGAGAAGTGC 234  
 Db 247 TGCAGTAAATACCTCGAAGATGCTTATGCACTAAGCCGTGTGCACTGAGAG----- 300  
 QY 235 ACTAAACAGAGAGCTGAAATTTTGGCTGAGAGCAAAACCTTGGCTGACGCTTGTCT 294  
 Db 301 -----GCACAGCAACTTGGCTAAGCGAGTAAGACTATGAGCTTGGCT 351  
 QY 295 GAGAGAGATATGGA 311  
 Db 352 GAGAGAGATATGATGA 368

RESULT 10  
 AF112443 506 bp mRNA linear PLN 10-DEC-1999  
 LOCUS Capsicum annuum thionin-like protein (Pepthi) mRNA, complete cds.  
 DEFINITION  
 ACCESSION AF112443  
 VERSION AF112443.1 GI:6552501  
 KEYWORDS  
 SOURCE Capsicum annuum  
 ORGANISM Capsicum annuum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Solanales; Solanaceae; Capsicum.  
 REFERENCE 1 (bases 1 to 506)  
 AUTHORS Oh,B.J., Ko,M.K., Kostenyuk,I., Shin,B. and Kim,K.S.  
 TITLE Coexpression of a defense gene and a thionin-like via different signal transduction pathways in pepper and Colletotrichum gloeosporioides interactions  
 JOURNAL Plant Mol. Biol. 41 (3), 313-319 (1999)  
 MEDLINE 2064966  
 PUBMED 10598099  
 REFERENCE 2 (bases 1 to 506)  
 AUTHORS Oh,B.J., Ko,M.K., Kostenyuk,I.A., Shin,B.C. and Kim,K.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-DEC-1998) Kumho life and Environmental Science Laboratory, Kumho Petrochemical Co., LTD., 1 Oryong-dong, Puk-gu, Kwangju 500-712, Korea  
 FEATURES  
 source location/Qualifiers  
 1..506  
 /organism="Capsicum annuum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4072"  
 gene 1..506  
 /gene="Pepthi"  
 CDS 10..264  
 /gene="Pepthi"  
 /codon\_start=1  
 /product="thionin-like protein"  
 /protein\_id="AAFI6413.1"  
 /db\_xref="GI:6552502"  
 /translation="MARSIYPMALFVLAATVFAVVGQKEICKELTPVCSDDL  
 CQKLCMEKEDHGHCFTILSKLCMKRCNATLATBELLA"

BASE COUNT 173 a 75 c 96 g 162 t  
 ORIGIN

Query Match 17.7%; Score 95.6; DB 8; Length 506;  
 Best Local Similarity 56.9%; Pred.No.7.3e-10;  
 Matches 312; Conservative 0; Mismatches 174; Indels 62; Gaps 5;

QY 1 ATGGCTGCTCTTGTGCTTCAATGCAATTTGGCAAGAGTCTTTGTGCC 60  
 Db 10 ATGGCTGCTTCAATTTCAATGCAATTTGCTTGGCAATGACCTCTTTGTGCT 69  
 QY 61 TATGAGTCAAGCTGAGATGCAAAAGAAAGACACATTTCTCGAATATGATT 120  
 Db 70 TATGGGAGTCAAGGCAAGAAATTTGCTTAAAGAGCTACAAACCTGTTAAAGTTCT 129  
 QY 121 ACCAAACACCAATGCAAGAAAGCTTGAT--CACTGAGAAATTTACTGATGCTATTGT 177

Db 130 AGTACCCCTCTATGTCAAAAACCTGTGTATGAGAGAGAAATATGAAATGTCATTGT 189  
 QY 178 AGCAAAATCTGTAGAGAGTGCCTATGTACTAAGCATGTGTGTTGATGAGAATGACT 237  
 Db 190 TTACATCTTCAAGCAAGTGTCTTATGCATGAAAGATGTA----- 229  
 QY 238 AAAACAGAGCTGAAAATTTTGGCTGAGAGCAAAAACCTTGGCTGACCTTGTCTGAA 297  
 Db 230 -----ATGCTAAACACCTCTGCAACGAAATGCTTGTCT 261  
 QY 298 GAGAGATATGATATGATATGATATGATGAGAAATTAAGATGACATGACAT 357  
 Db 262 TGA--AGAGACTTCTTCAAGCAGATATTAAGTTATGATTTAGTGTACACAA 318  
 QY 358 AATAAGT-----TTCTACCTTTCTTAAAGTGTAGCTATGTTGTTAAATGGCTTT 413  
 Db 319 AATTAATTAAGTGTGCTTCTTCTTAAAGGTAATTAATGTTGTTGTTGTTGTTGTT 378  
 QY 414 TATGACCTTTTATACCTTTAAATAGTGTGACCTTCATCCTTGTGCAATCTTGC 473  
 Db 379 TATGACCAATTTGACCAATTAATTAAGTTGTGACACATCAATCCTTCAATGATTTCT 438  
 QY 474 ACTAAGTTATTTGTGTACTTTTAAATGAAATGACCTTCTATGCTTTGTTGTTAAAAAA 533  
 Db 439 ATTAAG-----TTGTGTGTTTATGAAAGAGATGTTTACGCTTAAAAAAAAA 494  
 QY 534 AAAAAAA 541  
 Db 495 AAAAAAA 502

RESULT 11  
 AF112869 548 bp mRNA linear PLN 19-DEC-1999  
 LOCUS Capsicum annuum gamma-thionin 1 precursor (thionin) mRNA, complete cds.  
 DEFINITION  
 ACCESSION AF112869  
 VERSION AF112869.1 GI:6601330  
 KEYWORDS  
 SOURCE Capsicum annuum  
 ORGANISM Capsicum annuum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamids; Solanales; Solanaceae; Capsicum.  
 REFERENCE 1 (bases 1 to 548)  
 AUTHORS Hwang,B.K., Lee,S.C., Kim,Y.J. and Hong,J.K.  
 TITLE Molecular cloning and pathogen-induced expression of a thionin gene in pepper plants  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 548)  
 AUTHORS Hwang,B.K., Lee,S.C., Kim,Y.J. and Hong,J.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-DEC-1998) Dept. of Agricultural Biology, Korea University, Anam-dong, Sunbuk-gu, Seoul 136-701, Korea  
 FEATURES  
 source location/Qualifiers  
 1..548  
 /organism="Capsicum annuum"  
 /mol\_type="mRNA"  
 /cullivar="Handyul"  
 /db\_xref="taxon:4072"  
 /cissue\_type="leaf"  
 gene 1..548  
 /gene="thionin"  
 CDS 52..306  
 /gene="thionin"  
 /codon\_start=1  
 /product="gamma-thionin 1 precursor"  
 /protein\_id="AF18936.1"  
 /db\_xref="GI:6601331"  
 /translation="MARSIYPMALFVLAATVFAVVGQKEICKELTPVCSDDL  
 CQKLCMEKEDHGHCFTILSKLCMKRCNATLATBELLA"

sig\_peptide 52..114





[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2004, 08:46:51 ; Search time 14 Seconds

(without alignments)  
352.700 Million cell updates/sec

Title: US-10-072-809b-18

Sequence: 1 MARSLCFMFAFLARMLFVA.....LAERAKTLAALLERETMDN 105

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	95.8	105	1	THGF_TOBAC
2	468.5	84.9	106	1	THG1_NICPA
3	144	26.1	78	1	THG1_FERTIN
4	143.5	26.0	73	1	THG4_ARATH
5	138.5	25.1	74	1	P322_SOLTU
6	137.5	24.9	77	1	THG1_ARATH
7	137.5	24.9	77	1	THG3_ARATH
8	137	24.8	78	1	THGF_HELAN
9	135	24.5	74	1	DER2_CAPAN
10	132	23.9	52	1	DER2_SPIOL
11	127.5	23.1	77	1	THG2_ARATH
12	123	22.3	75	1	DEF1_CAPAN
13	112	20.3	80	1	APF2_ARATH
14	111	20.1	74	1	SRP_SOYBN
15	110	19.9	80	1	APF1_BRANA
16	108	19.6	80	1	APF4_RAPSA
17	107	19.4	47	1	THG1_MAIZE
18	106.5	19.3	79	1	THG1_RAPSA
19	105	19.0	80	1	APF3_ARATH
20	103	18.7	47	1	SA21_SORBI
21	103	18.7	47	1	SA13_SORBI
22	103	18.7	80	1	APF1_ARATH
23	103	18.7	80	1	APF2_RAPSA
24	102.5	18.6	79	1	APF3_ARANA
25	102	18.5	47	1	THG2_VIGNA
26	101	18.3	47	1	FAB2_VICFA
27	101	18.3	47	1	FAB2_VICFA
28	99	17.9	47	1	THG_HORVU
29	99	17.9	48	1	SI22_SORBI
30	97	17.6	80	1	APF4_ARATH
31	94.5	17.1	74	1	DR39_PEA
32	94	17.0	47	1	THG2_WHEAT
33	92.5	16.8	75	1	10KD_VIGNA

## ALIGNMENTS

RESULT 1	ID	THGF_TOBAC	STANDARD	PRT	105 AA.
AC	P32026	THGF_TOBAC			
DT	01-JUL-1993	(Rel. 26, Created)			
DT	01-JUL-1993	(Rel. 26, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE		Flower-specific gamma-thionin precursor.			
GN		FST.			
OS		Nicotiana glauca (Common tobacco).			
OC		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OX		Assteridae; lamiales; Solanales; Solanaceae; Nicotiana.			
OX		NCBI_TaxID=4097;			
RP		SEQUENCE FROM N.A.			
RC		SHRIMP-CV. M38. TISSUE=flower buds;			
RX		MEDLINE=92357021; PubMed=1495489;			
RA		Gu Q., Kawata E.E., Morse M.-J., Wu H.-M., Cheung A.Y.;			
RT		"A flower-specific cDNA encoding a novel thionin in tobacco.";			
RL		Mol. Gen. Genet. 234:89-96(1992).			
CC		- FUNCTION: INVOLVED IN FLORAL ORGANOGENESIS. MAY PLAY A PROTECTIVE			
CC		ROLE IN FLOWERS BY PROTECTING THE REPRODUCTIVE ORGANS FROM			
CC		POTENTIAL PATHOGEN ATTACK.			
CC		- SUBCELLULAR LOCATION: POSSIBLY THE CELL WALL OR VACUOLE.			
CC		- TISSUE SPECIFICITY: FLOWER. FOUND IN PETALS, STAMEN AND PISTILS,			
CC		BUT NOT IN SEEDS. IN PARTICULAR, ACCUMULATION IN A CONFIGURATION			
CC		SURROUNDING THE INNER REPRODUCTIVE WHORLS.			
CC		- DEVELOPMENTAL STAGE: ACCUMULATES IN DEVELOPING FLOWERS AND ITS			
CC		LEVEL DROPS AS FLOWERS MATURE.			
CC		- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.			
CC		-----			
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC		the European Bioinformatics Institute. There are no restrictions on its			
CC		use by non-profit institutions as long as its content is in no way			
CC		modified and this statement is not removed. Usage by and for commercial			
CC		entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC		or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC		-----			
CC		EMBL; Z11748; CAA77806.1; -			
CC		PIR; S23574; S23574.			
CC		DR InterPro; IPR002118; Gamma-thionin.			
CC		DR InterPro; IPR003614; Kncn1.			
CC		PFam; PF00304; Gamma-thionin; 1.			
CC		DR PRODOM; PD002594; G_Purothionin; 1.			
CC		SMART; SM00505; Kncn1; 1.			
CC		DR PROSITE; PS00940; GAMMA_THIONIN; 1.			
CC		KW Plant defense; Plant toxin; Signal.			
CC		FT CHAIN 1 25 POTENTIAL.			
CC		FT DISUFID 26 105 FLOWER-SPECIFIC GAMMA-THIONIN.			
CC		FT DISUFID 28 72 BY SIMILARITY.			
CC		FT DISUFID 39 59 BY SIMILARITY.			
CC		FT DISUFID 45 66 BY SIMILARITY.			
CC		FT DISUFID 49 68 BY SIMILARITY.			
CC		SEQUENCE 105 AA; 11750 MW; DAYF4511F0651BC3 CRC64;			

34	88.5	16.0	51	1	AF2A_SINAL	P30332 sinapis alb
35	88	15.9	47	1	THG1_WHEAT	P20158 triticum ae
36	86.5	15.7	72	1	D230_PEA	O01783 pisum sativ
37	85.5	15.5	46	1	AX2_BETU	P82010 beta vulgar
38	85.5	15.5	51	1	APF1_SINAL	P30231 sinapis alb
39	84.5	15.3	46	1	AX1_BETU	P81493 beta vulgar
40	78	14.1	47	1	SI21_SORBI	P21923 sorghum bic
41	77	13.9	70	1	DMYC_DROME	P41964 drosophila
42	76	13.8	47	1	PSD2_PEA	P81930 pisum sativ
43	76	13.8	47	1	THG2_MAIZE	P81009 zea mays (m
44	73.5	13.3	297	1	SGS4_DROME	O00725 drosophila
45	68	12.3	332	1	FSA_BRARE	O9yiv4 brachydanio

Query Match 95.8%; Score 529; DB 1; Length 105;  
 Best Local Similarity 95.2%; Pred. No. 2.6e-49;  
 Matches 100; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MARSLCFMAFALLIARMLFVAYEVQARCKTESNTPPGICITKPPCKRACISEKFTDGHCS 60  
 DB 1 MARSLCFMAFALLIARMLFVAYEVQARCKTESNTPPGICITKPPCKRACISEKFTDGHCS 60

QY 61 KIIRRCICTKPCVDEKMTKTGAETLAEAKTAAALIEEIMDN 105  
 DB 61 KIIRRCICTKPCVDEKMTKTGAETLAEAKTAAALIEEIMDN 105

RESULT 2  
 THG1\_NICPA STANDARD; PRT; 106 AA.  
 AC 024115;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gamma-chionin 1 precursor.  
 OS Nicotiana glauca.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.  
 CX NCBI\_TaxID=62141;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Leaf;  
 RA Komori T., Yamada S., Imaseki H.;  
 RT "A cDNA clone for gamma-chionin from Nicotiana glauca";  
 RL (In) Plant Gene Register PGR97-132.  
 CC -1- INDUCTION: By salt stress.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AB005250; BAA21325.1;  
 DR InterPro: IPR002118; Gamma-chionin.  
 DR InterPro: IPR003614; Knott1.  
 DR Pfam: PF00304; Gamma-chionin; 1.  
 DR ProDom: PD002594; G\_Purochionin; 1.  
 DR SMART: SM00505; Knott1; 1.  
 DR PROSITE: PS00940; GAMMA\_THIONIN; FALSE\_NEG.  
 KW plant defense; Plant toxin; Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 106 GAMMA-THIONIN 1.  
 FT DISULFID 29 73 BY SIMILARITY.  
 FT DISULFID 40 60 BY SIMILARITY.  
 FT DISULFID 46 67 BY SIMILARITY.  
 FT DISULFID 50 69 BY SIMILARITY.  
 FT DISULFID 50 69 BY SIMILARITY.  
 SQ SEQUENCE 106 AA; 11722 MW; 269850E116D775A4 CRC64;

Query Match 84.9%; Score 468.5; DB 1; Length 106;  
 Best Local Similarity 84.0%; Pred. No. 6.7e-43;  
 Matches 89; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 MARSLCFMAFALLIARMLFVAYEVQARCKTESNTPPGICITKPPCKRACISEKFTDGHCS 59  
 DB 1 MARSLCFMAFALLIARMLFVAYEVQARCKTESNTPPGICITKPPCKRACISEKFTDGHCS 60

QY 60 SKILRRCLCTKPCVDEKMTKTGAETLAEAKTAAALIEEIMDN 105  
 DB 61 SKILRRCLCTKPCVDEKMTKTGAETLAEAKTAAALIEEIMDN 106

RESULT 3  
 THG1\_PETIN STANDARD; PRT; 78 AA.  
 AC Q40901;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gamma-chionin homolog PPT precursor.  
 OS Petunia integrifolia (violet-flowered petunia) (Petunia inflata).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; lamids; Solanales; Solanaceae; Petunia.  
 CX NCBI\_TaxID=4103;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Pistil;  
 RX MEDLINE=95036017; PubMed=7948892;  
 RA Karunahandaa B., Singh A., Kao T.H.;  
 RT "Characterization of a predominantly pistil-expressed gene encoding a  
 RT gamma-chionin-like protein of Petunia inflata";  
 RL Plant Mol. Biol. 26:459-464(1994).  
 CC -1- FUNCTION: May be involved in the defense of the pistil against  
 CC pathogen infection.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE PISTIL DURING  
 CC ALL STAGES OF FLOWER DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; L21713; AAA64740.1;  
 DR PIR: S52634; S52634.  
 DR HSRP: P20230; IGPT.  
 DR InterPro: IPR002118; Gamma-chionin.  
 DR InterPro: IPR003614; Knott1.  
 DR Pfam: PF00304; Gamma-chionin; 1.  
 DR ProDom: PD002594; G\_Purochionin; 1.  
 DR SMART: SM00505; Knott1; 1.  
 DR PROSITE: PS00940; GAMMA\_THIONIN; 1.  
 KW plant defense; Signal.  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 78 GAMMA-THIONIN HOMOLOG PPT.  
 FT DISULFID 34 78 BY SIMILARITY.  
 FT DISULFID 45 65 BY SIMILARITY.  
 FT DISULFID 51 72 BY SIMILARITY.  
 FT DISULFID 55 74 BY SIMILARITY.  
 SQ SEQUENCE 78 AA; 8799 MW; 263DB3AC9D619265 CRC64;

Query Match 26.1%; Score 144; DB 1; Length 78;  
 Best Local Similarity 35.9%; Pred. No. 1.2e-08;  
 Matches 28; Conservative 14; Mismatches 30; Indels 6; Gaps 1;

QY 1 MARSLCFMAFALLIARMLFVAYEVQARCKTESNTPPGICITKPPCKRACISEKFTDGHCS 54  
 DB 1 MARSLCFMAFALLIARMLFVAYEVQARCKTESNTPPGICITKPPCKRACISEKFTDGHCS 60

QY 55 TDGHCCKIIRRCICTKPC 72  
 DB 61 IGCNCRAPRRRCPTNRC 78

RESULT 4  
 THG4\_ARATH STANDARD; PRT; 73 AA.  
 AC Q92018;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)



DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gamma-thionin homolog At2g02140 precursor.  
 GN AT2g02140 OR F504.9.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDLINE=2008487; PubMed=10617197;  
 RA Lin X., Xau S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buel C.R., Ketchum K.A., Lee J.J., Rinning C.M., Koo H.L.,  
 RA Mottet K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,  
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,  
 RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,  
 RA Niernmen W.C., White O., Eisen J.A., Salzberg S.L., Frazer C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 402:761-768(1999).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AC005936; AAC97220.1; -  
 DR PIR; D84433; D84433.  
 DR HSSP; P20230; 1GPT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knt1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; G\_Purothionin; 1.  
 DR SMART; SM00505; Knt1; 1.  
 DR PROSITE; PS00940; GAMMA-THIONIN; 1.  
 KW Plant defense; Signal; Multigene family.  
 FT SIGNAL 1 28  
 FT CHAIN 1 73  
 FT DISULFID 31 72  
 FT DISULFID 42 62  
 FT DISULFID 48 67  
 FT DISULFID 52 69  
 FT DISULFID 52 69  
 SQ SEQUENCE 73 AA; 7718 MW; DA27D295AC3739DA CRC64;  
 Query Match 26.0%; Score 143.5; DB 1; Length 73;  
 Best Local Similarity 40.0%; Pred. No. 1.2e-08;  
 Matches 30; Conservative 13; Mismatches 27; Indels 5; Gaps 2;  
 QY 1 MARSICFWAPAILARMLFVAY---EVARBECKTSNTPPGICITPKPCRAKISSEKTDG 57  
 DB 1 MKLSIRLISALIMSVMLFATGMPVEARCEPSNKFQGVCLNSGCAACAPSEBGS 60  
 QY 58 HCSKILRRCLCTKPC 72  
 DB 61 RCSSL-RCYCSKAC 73  
 RESULT 5  
 ID P322 SOLTU STANDARD; PRT; 74 AA.  
 AC P20346;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable protease inhibitor P322 precursor.

OS Solanum tuberosum (Potato).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.  
 CC NCBI\_TaxID=4113;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Bintje;  
 RA Stiekema W.J., Heidekamp F., Dirkee W.G., van Beckum J., de Haan P.,  
 RA ten Bosch C., Louwerse J.D.;  
 RT "Molecular cloning and analysis of four potato tuber mRNAs.";  
 RL Plant Mol. Biol. 11:255-269(1988).  
 CC -1- TISSUE SPECIFICITY: TUBER.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X13180; CAA31577.1; -  
 DR PIR; S05594; S05594.  
 DR HSSP; P20230; 1GPT.  
 DR InterPro; IPR002118; Gamma-thionin.  
 DR InterPro; IPR003614; Knt1.  
 DR Pfam; PF00304; Gamma-thionin; 1.  
 DR ProDom; PD002594; G\_Purothionin; 1.  
 DR SMART; SM00505; Knt1; 1.  
 DR PROSITE; PS00940; GAMMA-THIONIN; 1.  
 KW Serine protease inhibitor; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 1 74  
 FT DISULFID 30 74  
 FT DISULFID 41 61  
 FT DISULFID 47 68  
 FT DISULFID 51 70  
 FT DISULFID 51 70  
 SQ SEQUENCE 74 AA; 8414 MW; FAFEL87038F4B5A CRC64;  
 Query Match 25.1%; Score 138.5; DB 1; Length 74;  
 Best Local Similarity 40.3%; Pred. No. 4.2e-08;  
 Matches 29; Conservative 11; Mismatches 25; Indels 7; Gaps 2;  
 QY 7 FAPAILARMLFVAYEV-----QARBECKTSNTPPGICITPKPCRAKISSEKTDGCS 60  
 DB 4 FAFIFLLA-MLVYARKMPRIAEKRHCESHRTKPCCTDSNCAVCETERRSGNCH 62  
 QY 61 KILRRCLCTKPC 72  
 DB 63 GFRRCFCYKPC 74  
 RESULT 6  
 ID THG1 ARATH STANDARD; PRT; 77 AA.  
 AC Q39182; Q42011;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gamma-thionin homolog At2g02100 precursor.  
 GN AT2G02100 OR F504.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. C24; TISSUE=flower buds;  
 RA Yu D.Y., Quidley F., Maché R.;  
 RT "Isolation and expression of a cDNA encoding protease inhibitor



OS Helianthus annuus (Common sunflower).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Asterales; Asteraceae; Asteroidae;  
 OC Heliantheae; Helianthus.  
 OK NCBI\_TaxID=4232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=flower;  
 RA Urdangarin M.C., de la Canal L.;  
 RT "A defensin gene expressed in sunflower inflorescence."  
 RL Plant Physiol. Biochem. 38:253-258(2000).  
 CC -1- FUNCTION: MAY PLAY A PROTECTIVE ROLE IN FLOWERS BY PROTECTING THE  
 CC REPRODUCTIVE ORGANS FROM POTENTIAL PATHOGEN ATTACK.  
 CC -1- SUBCELLULAR LOCATION: CELL WALL OR VACUOLE (POTENTIAL).  
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN FLOWERS AND TO A LESSER  
 CC EXTENT IN LEAVES. LOWER LEVELS IN HYPOCOTYLS. NO EXPRESSION IN  
 CC ROOTS AND COTYLEDONS.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED PROGRESSIVELY DURING FLOWER  
 CC DEVELOPMENT REACHING THE HIGHEST LEVEL IN THE MATURE FERTILIZED  
 CC FLOWER STAGE.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF178634; AAF72042.1; -  
 DR EMBL: AF141131; AAF6591.1; -  
 DR HSSP: P20230; IGPT.  
 DR InterPro: IPR002118; Gamma-thionin.  
 DR InterPro: IPR003614; Knoc1.  
 DR Pfam: PF00304; Gamma-thionin; 1.  
 DR ProDom: PD002594; G\_Purothionin; 1.  
 DR SMART: SM00505; Knoc1; 1.  
 DR PROSITE: PS00940; GAMMA\_THIONIN; 1.  
 DR Plant defense; Signal.  
 KW SIGNAL  
 FT CHAIN 1 20  
 FT DISULFID 21 78  
 FT DISULFID 34 78  
 FT DISULFID 45 65  
 FT DISULFID 51 72  
 FT DISULFID 55 74  
 FT DISULFID 55 74  
 SQ SEQUENCE 78 AA; 8627 MW; 1AD6A98582149A2D CRC64;  
 Query Match 24.8%; Score 137; DB 1; Length 78;  
 Best Local Similarity 35.9%; Pred. No. 6,4e-08;  
 Matches 28; Conservative 12; Mismatches 32; Indels 6; Gaps 1;  
 QY 1 MARSLCPMAFALRLFLFAVE-----VQARECKTESNTPFGICITPCKKACISKEF 54  
 Db 1 MMSMMFALALVWVCLANENGELVEARTCSQSHKFGKGLSDTNCANVCHSERF 60  
 QY 55 TDGHCCKILRCLCTKPC 72  
 Db 61 SGGKCGFRRCCTTHC 78  
 RESULT 9  
 DEF2 CAPAN STANDARD; PRT; 74 AA.  
 ID DEF2 CAPAN  
 AC 065740;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Defensin J1-2 precursor.  
 OS Capsicum annuum (Bell pepper).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamids; Solanales; Solanaceae; Capsicum.  
 OK NCBI\_TaxID=4072;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=cv. Yolo Wonder; TISSUE=fruit;  
 RX MEDLINE=97037730; Pubmed=8883377;  
 RA Meyer B., Houine G., Forzetta-Romero J., Schantz M.L., Schantz R.;  
 RT "Fruit-specific expression of a defensin-type gene family in bell  
 RT pepper. Upregulation during ripening and upon wounding."  
 RL Plant Physiol. 112:615-622(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=cv. Yolo Wonder; TISSUE=fruit;  
 RX MEDLINE=9905242; Pubmed=9790581;  
 RA Houine G., Meyer B., Schantz R.;  
 RT "Alteration of the expression of a plant defensin gene by exon  
 RT splicing in bell pepper (Capsicum annuum L.)."  
 RL Mol. Gen. Genet. 259:504-510(1998).  
 CC -1- FUNCTION: PLANT DEFENSE PEPTIDE WITH ANTIFUNGAL ACTIVITY AGAINST  
 CC FOXSPORUM AND B.CINEREA.  
 CC -1- SUBUNIT: Monomer (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.  
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIPENING.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X95730; CA65046.1; -  
 DR HSSP: P20230; IGPT.  
 DR InterPro: IPR002118; Gamma-thionin.  
 DR InterPro: IPR003614; Knoc1.  
 DR Pfam: PF00304; Gamma-thionin; 1.  
 DR ProDom: PD002594; G\_Purothionin; 1.  
 DR SMART: SM00505; Knoc1; 1.  
 DR PROSITE: PS00940; GAMMA\_THIONIN; 1.  
 DR Plant defense; Fungicide; Signal.  
 KW SIGNAL  
 FT CHAIN 1 27  
 FT DISULFID 28 74  
 FT DISULFID 30 74  
 FT DISULFID 41 61  
 FT DISULFID 47 68  
 FT DISULFID 51 70  
 FT DISULFID 51 70  
 SQ SEQUENCE 74 AA; 8249 MW; D92D8F06F39E1552 CRC64;  
 Query Match 24.5%; Score 135; DB 1; Length 74;  
 Best Local Similarity 34.9%; Pred. No. 9,9e-08;  
 Matches 22; Conservative 12; Mismatches 29; Indels 0; Gaps 0;  
 QY 10 FALRLFLFAVEVQARECKTESNTPFGICITPCKKACISKEFDGHCCKILRCLCT 69  
 Db 12 FLMMFLVATGMVAEARTCSQSHKFGKGLFSKNCVCHTGFNGHCGRFRRCFCT 71  
 QY 70 KPC 72  
 Db 72 RHC 74  
 RESULT 10  
 DEF2 SPIOL STANDARD; PRT; 52 AA.  
 ID DEF2 SPIOL  
 AC 81571;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Antimicrobial peptide D2 (So-D2) (Defensin D2) (Fragment).  
 OS Spinacia oleracea (Spinach).

```

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RL thaliana.",
CC Nature 402:761-768(1999).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
-----
CC
DR EMBL; Z17665; CAI79029.1; -.
DR PIR; B84433; B84433.
DR HSP; P20230; IGPT.
DR InterPro; IPR002118; Gamma-thionin.
DR InterPro; IPR003614; Knott1.
DR Pfam; PF00304; Gamma-thionin; 1.
DR ProDom; PD002594; G_Purothionin; 1.
DR SMART; SM00505; Knott1; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
DR Plant defense; signal; Multigene family.
FT SIGNAL 1 30
FT CHAIN 1 77
FT DISULFID 33 77
FT DISULFID 44 64
FT DISULFID 50 71
FT DISULFID 54 73
FT CONFLICT 19 19
FT CONFLICT 69 69
SQ SEQUENCE 77 AA; 8578 MW; FBTB5AF727A945 CRC64;
Query Match 23.1%; Score 127.5; DB 1; Length 77;
Best Local Similarity 33.8%; Pred. No; 6.4e-07;
Matches 26; Conservative 12; Mismatches 34; Indels 5; Gaps 1.
OY 1 MARSLCFMAFPIIARMLFVA-----YEQVAECTESTSPGICITKPPCKRACISEFT 55
DB 1 MKFEMRIISAVLFLVMIFVATGMPYVTEATCASQSGRFKACVSDTNCVENCNBGF 60
OY 56 DGHCSKTLRCLCTKPC 72
DB 61 GGDGRGPRRCFCITRNC 77
RESULT 12
DEFL_CAPAN STANDARD; PRT; 75 AA.
AC Q43413;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Defensin j1-1 precursor.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_Taxid=4072;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 28-69, AND CHARACTERIZATION.
RC STRAIN=cv. Yolo Wonder; TISSUE=fruit;
RX MEDLINE=97037730; PubMed=8883377;
RA Meyer B., Houline G., Pozueta-Romero J., Schantz M.L., Schantz R.;
RT "Fruit-specific expression of a defensin-type gene family in bell
RT pepper. Upregulation during ripening and upon wounding.";
RL Plant Physiol. 112:615-622(1996).
CC -1- FUNCTION: PLANT DEFENSE PEPTIDE WITH ANTIFUNGAL ACTIVITY AGAINST
CC F. OXYSPORUM AND B. CINERA.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Secreted.

```

```

RT   thaliana.";
RL   Nature 402:761-768(1999).
CC   -!- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC
CC   CATIONS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PLANT DEENSGIN FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for Commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AC004747; AACJ1234.1; -.
DR   PIR; T02622; T02622.
DR   HSSP; P30231; IAYU.
DR   InterPro; IPR002118; Gamma-thionin.
DR   InterPro; IPR003614; Knott1.
DR   Pfam; PF00304; Gamma-thionin; 1.
DR   ProDom; PD002594; G-Purothionin; 1.
DR   SMART; SM00505; Knott1; 1.
DR   PROSITE; PS00940; GAMMA_THIONIN; 1.
KW   Plant defense; Fungicide; Signal; Multigene family;
KW   Pyroglutidone carboxylic acid.
FT   SIGNAL          1..29
FT                     BY SIMILARITY.
FT   CHAIN           30..80
FT                     PROBABLE CYSTEINE-RICH ANTIFUNGAL PROTEIN
FT                                     ATG262610.
FT   MOD_RES         30..30
FT                       PYROGLUTIDONE CARBOXYLIC ACID (BY
FT                           SIMILARITY).
FT   DISULFID        33..80
FT                       BY SIMILARITY.
FT   FT             44..65
FT                       BY SIMILARITY.
FT   DISULFID       50..74
FT                       BY SIMILARITY.
FT   DISULFID       54..76
FT                       BY SIMILARITY.
SQ   SEQUENCE      80 AA;  8580 MW;  99EBE0D443AD67B CRC64;
Query Match              20.3%; Score 112; DB 1; Length 80;
Best Local Similarity    32.5%; Pred. No. 2.9e-05;
Matches 26; Conservative 16; Mismatches 30; Indels 8; Gaps 4.
QY      1 MARSLCEFAPIFLRMLEFVAVE---VGARE-CKTESNTFGPICITKPCRFACIS-EKF 54
DB      1 MAKRAAITTFPALVLPFAPEAPATWQAQLCEKPSGTWAGVCNSNACNQCLINLEGA 60
QY      55 TDGHCSKTL--RCLCTKPC 72
DB      61 KHGSCHVFPAHKCIQYPFC 80
RESULT 14
SRP_SOYBN STANDARD; PROT; 74 AA.
AC Q07502;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE 8.4 kDa sulfur-rich protein precursor (SB60 protein) (Probable
DE     protease inhibitor PJ322).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
CX NCBI_TaxID=3847;
[!]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Palda1; TISSUE=seed;
RX MEDLINE=94105312; PubMed=8278516;
RA Choi Y., Choi Y.D., Lee J.S.;
RT "Nucleotide sequence of a cDNA encoding a low molecular weight
RT     sulfur-rich protein in soybean seeds.";
CC Plant Physiol. 101:699-700(1993).
CC -!- FUNCTION: MAY FUNCTION AS A SULFUR-RICH STORAGE PROTEIN OR A

```

PROTEINASE INHIBITOR.  
 -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: Z13956; CAA78359.1; -  
 DR PIR: S24965; S24965.  
 DR HSSP: P20230; 1GPT.  
 DR InterPro: IPR002118; Gamma-thionin.  
 DR InterPro: IPR003614; Knott.  
 DR Pfam: PF00304; Gamma-thionin; 1.  
 DR ProDom: PD002594; G\_purothionin; 1.  
 DR SMART: SM00505; Knott; 1.  
 DR PROSITE: PS00940; GAMMA\_THIONIN; 1.  
 KW Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 74 8.4 kDa SULFUR-RICH PROTEIN.  
 FT DISULFID 31 74 BY SIMILARITY.  
 FT DISULFID 42 62 BY SIMILARITY.  
 FT DISULFID 48 68 BY SIMILARITY.  
 FT DISULFID 52 70 BY SIMILARITY.  
 SQ SEQUENCE 74 AA; 8437 MW; 38E93064D615CE8 CRC64;  
 Query Match 20.1%; Score 111; DB 1; Length 74;  
 Best Local Similarity 34.7%; Pred. No. 3.5e-05;  
 Matches 26; Conservative 14; Mismatches 29; Indels 6; Gaps 3;  
 QY 1 MARSLCFMAFALLAMLP---VAYVQARECKTSNFPGICITKPCRCACISAKFTG 57  
 DB 3 MKRSCGF--FFLLLVFASGVVQTEGVCSQSHGHGLCDNDHNCALVCEGSG 60  
 QY 58 HCSKILRCLCTKPC 72  
 DB 61 RC-KSRRCFCRTRIC 74  
 RESULT 15  
 AFPI BRANA STANDARD; PRT; 80 AA;  
 ID AFPI BRANA STANDARD; PRT; 80 AA;  
 AC P30225; Q41163;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cysteine-rich antifungal protein 1 precursor (AFPI).  
 GN AFPI.  
 OS Brassica napus (Rape), and  
 OS Raphanus sativus (Radish).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Brassica.  
 CX NCBI\_Taxid=3708, 3726;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Brassica napus; STRAIN=cv. Ronde Rode Kleine Wiltpunt; TISSUE=Seed;  
 RC MEDLINE=9529350; PubMed=780308;  
 RA Terias F.R.G., Eggelmont K., Kovaleva V., Raikhel N.V., Osborn R.W.,  
 RA Kester A., Rees S.B., Torrekens S., Van Leuven F., Vanderleyden J.,  
 RA Cammue B.P.A., Broekaert W.F.;  
 RT "Small cysteine-rich antifungal proteins from radish: their role in  
 RT host defense."  
 RT Plant Cell 7:568-573 (1995).  
 RL [2]  
 RP SEQUENCE OF 30-73.  
 RC SPECIES=Brassica napus; TISSUE=Seed;  
 RC MEDLINE=9318130; PubMed=8422949;  
 RA Terias F.R.G., Torrekens S., Van Leuven F., Osborn R.W.,  
 RA Vanderleyden J., Cammue B.P.A., Broekaert W.F.;

RT "A new family of basic cysteine-rich plant antifungal proteins from  
 RT Brassicaceae species.";  
 RL FEBS Lett. 316:233-240 (1993).  
 RN [3]  
 RP SEQUENCE OF 30-73. TISSUE=Seed;  
 RC SPECIES=Brassica napus; TISSUE=Seed;  
 RC MEDLINE=92348373; PubMed=1639777;  
 RA Terias F.R.G., Schoofs H.M.E., de Bolle M.F.C., van Leuven F.,  
 RA Rees S.B., Vanderleyden J., Cammue B.P.A., Broekaert W.F.;  
 RT "Analysis of two novel classes of plant antifungal proteins from  
 RT radish (Raphanus sativus L.) seeds."  
 RL J. Biol. Chem. 267:15301-15309 (1992).  
 CC -1- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC  
 CC CATIONS.  
 CC -1- SUBUNIT: FORMS OLIGOMERS IN ITS NATIVE STATE.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: U18557; AAA69541.1; -  
 DR PIR: T10176; T10176.  
 DR HSSP: P30231; 1AYJ.  
 DR InterPro: IPR002118; Gamma-thionin.  
 DR InterPro: IPR003614; Knott.  
 DR Pfam: PF00304; Gamma-thionin; 1.  
 DR ProDom: PD002594; G\_purothionin; 1.  
 DR SMART: SM00505; Knott; 1.  
 DR PROSITE: PS00940; GAMMA\_THIONIN; 1.  
 KW Plant defense; Fungicide; Signal; Multigene family;  
 KW Pyroliidone carboxylic acid.  
 FT SIGNAL 1 29  
 FT CHAIN 30 80  
 FT MOD RES 30 80 CYS-THIONIN RICH ANTIFUNGAL PROTEIN 1.  
 FT DISULFID 33 80 PYROLIIDONE CARBOXYLIC ACID.  
 FT DISULFID 44 65 BY SIMILARITY.  
 FT DISULFID 50 74 BY SIMILARITY.  
 FT DISULFID 54 76 BY SIMILARITY.  
 SQ SEQUENCE 80 AA; 8734 MW; 05B90FA6C8D6C2B CRC64;  
 Query Match 19.9%; Score 110; DB 1; Length 80;  
 Best Local Similarity 32.5%; Pred. No. 4.8e-05;  
 Matches 26; Conservative 16; Mismatches 30; Indels 8; Gaps 4;  
 QY 1 MARSLCFMAFALLAMLPVAYVQARECKTSNFPGICITKPCRCACISAKFTG 54  
 DB 1 MAKRSALIALFALLVFAAFAPVFAKICERSGIVSGVCGNNCKRQCINLEBA 60  
 QY 55 TDGHCSTL--RCLCTKPC 72  
 DB 61 RHGSCNVPFAHKICITCYPC 80

Search completed: January 28, 2004, 08:47:20  
 Job time : 16 secs